

# STATE-INDEPENDENT IMPORTANCE SAMPLING FOR RANDOM WALKS WITH REGULARLY VARYING INCREMENTS

Karthiek R. A. Murthy <sup>\*a</sup>, Sandeep Juneja<sup>a</sup>, and Jose Blanchet<sup>b</sup>

<sup>a</sup>Tata Institute of Fundamental Research, Mumbai, India

<sup>b</sup>Columbia University, New York, United States

## Abstract

We develop state-independent importance sampling based efficient simulation techniques for two commonly encountered rare event probabilities associated with random walk ( $S_n : n \geq 0$ ) having i.i.d. regularly varying heavy-tailed increments; namely, the level crossing probabilities when the increments of  $S_n$  have a negative mean, and the large deviation probabilities  $\mathbb{P}\{S_n > b\}$ , as both  $n$  and  $b$  increase to infinity for the zero mean random walk. Exponential twisting based state-independent methods, which are effective in efficiently estimating these probabilities for light-tailed increments are not applicable when these are heavy-tailed. To address the latter case, more complex and elegant state-dependent efficient simulation algorithms have been developed in the literature over the last few years. We propose that by suitably decomposing these rare event probabilities into a dominant and further residual components, simpler state-independent importance sampling algorithms can be devised for each component resulting in composite unbiased estimators with a desirable vanishing relative error property. When the increments have infinite variance, there is an added complexity in estimating the level crossing probabilities as even the well known zero variance measures have an infinite expected termination time. We adapt our algorithms so that this expectation is finite while the estimators remain strongly efficient. Numerically, the proposed estimators perform at least as well, and sometimes substantially better than the existing state-dependent estimators in the literature.

## 1. INTRODUCTION

In this paper, we develop importance sampling algorithms involving simple, state-independent changes of measure for the efficient estimation of large deviation probabilities, and level crossing probabilities of random walks with regularly varying heavy-tailed increments. Specifically, let  $(X_n : n \geq 1)$  denote a sequence of zero mean independent and identically distributed (i.i.d.) random variables such that  $\mathbb{P}(X_n > x) = L(x)x^{-\alpha}$ , for some  $\alpha > 1$  and a slowly varying function<sup>1</sup>  $L(\cdot)$ . Note that  $\alpha > 2$  ensures finite variance for  $X_n$  whereas  $\alpha < 2$  implies that it has infinite variance. Set  $S_0 = 0$  and  $S_n = X_1 + \dots + X_n$ , for  $n \geq 1$ . Given  $\mu > 0$ , define  $M := \sup_n (S_n - n\mu)$ , and  $\tau_b := \inf\{n \geq 0 : S_n - n\mu > b\}$ . We are interested in the importance sampling based efficient estimation of:

1. Large deviation probabilities  $\mathbb{P}\{S_n > b\}$  for  $b > n^{\beta+\epsilon}$  with  $\beta := (\alpha \wedge 2)^{-1}$  as  $n \nearrow \infty$ , and
2. Level crossing probabilities  $\mathbb{P}\{\tau_b < \infty\}$ , or equivalently, the tail probabilities  $\mathbb{P}\{M > b\}$  as  $b \nearrow \infty$ .

---

<sup>\*</sup>Corresponding author. E-mail address: [kamurthy@tifr.res.in](mailto:kamurthy@tifr.res.in).

<sup>1</sup>That is,  $\lim_{x \rightarrow \infty} L(tx)/L(x) = 1$  for any  $t > 0$ ; prominent examples for slowly varying functions include  $(\log x)^\beta$  for any  $\beta \in \mathbb{R}$ .

For brevity, we refer to former as large deviations probabilities and the latter as level crossing probabilities. Our methodology for estimating the large deviations probabilities easily extends to the efficient estimation of  $\mathbb{P}\{S_N > u\}$  for a random  $N$ , when  $N$  is light-tailed<sup>2</sup> and independent of increments  $\{X_n\}$  (popular in literature are  $N$  fixed or geometrically distributed) as  $u \nearrow \infty$ . However, in the interest of space, we do not explicitly consider the ‘random sum tail probabilities’ estimation problem in this paper.

Importance sampling via appropriate change of measure has been extremely successful in efficiently simulating rare events, and has been studied extensively in both the light and heavy tailed settings (see, e.g., Asmussen and Glynn (2007) for an introduction to rare event simulation and applications). In importance sampling for random walks, state-dependence essentially means that the sampling distribution for generating the increment  $X_k$  depends on the realized values of  $X_1, \dots, X_{k-1}$  (typically, through  $S_{k-1}$ ); state-independence on the other hand implies that samples of  $X_1, \dots, X_n$  can be drawn independently. State-independent methods often enjoy advantages over state-dependent ones in terms of complexity of generating samples and ease of implementation. The zero-variance changes of measure for estimating the large deviations and the level crossing probabilities are well known and are state-dependent (see, e.g., Juneja and Shahabuddin 2006). While typically unimplementable, they provide guidance in search for implementable approximately zero variance importance sampling techniques.

In the light-tailed settings, large deviations analysis can be used to show that exponential twisting based state-independent importance sampling well approximates the zero variance measure (see, e.g., Asmussen and Glynn (2007)) and also efficiently estimates the large deviations as well as level crossing probabilities (see, e.g., Sadowsky and Bucklew (1990) and Siegmund (1976)). However, development of state independent techniques for these probabilities is harder in the heavy-tailed settings. Asmussen et al. (2000) provide an account of failure of simple large deviations based simulation methods that approximate zero-variance measure in heavy-tailed systems. Bassamboo et al. (2007) prove that any state-independent importance sampling change of measure cannot efficiently simulate level crossing probability in a busy cycle of a heavy tailed random walk. The fact that the zero-variance measures for estimating both the large deviations and the level crossing probabilities are state-dependent, and the above mentioned negative results, have motivated research over the last few years in development of complex and elegant state-dependent algorithms to efficiently estimate these probabilities (see, e.g., Dupuis et al. (2007), Blanchet and Glynn (2008), Blanchet and Liu (2008, 2012), and Chan et al. (2012)).

In this paper we introduce simple state-independent change of measures to estimate the large deviations and the level crossing probabilities with regularly varying increments. We show that the proposed methods are provably efficient<sup>3</sup> and perform at least as well as the existing state-dependent algorithms. Thus our key contribution is to question the prevailing view that one needs to resort to state-dependent methods for efficient computation of rare event probabilities involving ‘large number’ of heavy-tailed random variables. A key idea to be exploited in the estimation of probabilities considered is the fact that the corresponding rare event occurrence is governed by the “single big jump” principle, that is, the most likely paths leading to the occurrence of the rare event have one of the increments taking large value (see, for e.g., Foss et al. (2011) and the references therein). Our approach for estimating the large deviations

---

<sup>2</sup>As is well-known,  $X$  is light-tailed if the moment generating function  $\mathbb{E}[\exp(\theta X)]$  is finite for some  $\theta > 0$ , and is heavy-tailed otherwise.

<sup>3</sup>We show that the estimators have asymptotically vanishing relative error; this corresponds to their coefficient of variation converging to zero as the event becomes rarer. We also have a related weaker notion of strong efficiency where the coefficient of variation of the estimators, and subsequently the number of i.i.d. replications required, remains bounded as the event becomes rarer. Weak efficiency is another standard notion of performance in rare event simulation corresponding to a slow increase in the number of replications required as the event becomes rarer. These are briefly reviewed in Section 2.2.

probability  $\mathbb{P}\{S_n > b\}$  relies on decomposing it into a dominant and a residual component, and developing efficient estimation techniques for both. For estimating the level crossing probability  $\mathbb{P}\{\tau_b < \infty\}$ , in addition to such a decomposition, we partition the event of interest into several blocks that are sampled using appropriate randomization. When the increments have infinite variance, there is an added complexity in estimating the level crossing probabilities as even the well known zero variance measure is known to have an infinite expected termination time. We modify our algorithms so that this expectation remains finite while the estimator remains strongly efficient although it may no longer have asymptotically vanishing relative error.

Our specific contributions are as follows:

1. We provide importance sampling estimators that achieve asymptotically vanishing relative error for the estimation of  $\mathbb{P}\{S_n > b\}$ , as  $n \nearrow \infty$ . Given  $n$  and  $\epsilon > 0$ , our simulation methodology is uniformly efficient for values of  $b$  larger than  $n^{\frac{1}{2}+\epsilon}$  when the increments  $X_n$  have finite variance, and for  $b > n^{\frac{1}{\alpha}+\epsilon}$  in the case of increments having infinite variance – thus operating throughout the large deviations regime where the well-known asymptotics  $\mathbb{P}\{S_n > b\} \sim n\bar{F}(b)$  hold. Further, this is the first instance that we are aware of where efficient simulation techniques for the large deviations probability include the case of increments having infinite variance, which is not uncommon in practical applications involving heavy-tailed random variables.
2. For  $\alpha > 1$ , we develop unbiased estimators for level crossing probabilities  $\mathbb{P}\{\tau_b < \infty\}$  that achieve vanishing relative error as  $b \nearrow \infty$ . These estimators require an overall computational effort that scales as  $O(b)$  when variance of  $X_n$  is finite. This is similar to the complexity of the zero variance operator since, as is well known, the latter requires order  $\mathbb{E}[\tau_b | \tau_b < \infty]$  computation in generating a single sample and this is known to be linear in  $b$  when the variance of increments is finite (see Asmussen and Kluppelberg (1996)). However, since  $\mathbb{E}[\tau_b | \tau_b < \infty] = \infty$  for the case of increments having infinite variance, even the zero-variance measure (even if implementable) is no longer viable because from a computational standpoint, any useful estimator needs to have finite expected replication termination time. For random walks with infinite variance increments, we develop algorithms such that:
  - (a) For  $\alpha > 1.5$ , the associated estimators are strongly efficient and have  $O(b)$  expected termination time. As a converse, we also prove that for  $\alpha < 1.5$  no algorithm can be devised in our framework that has both the variance and expected termination time simultaneously finite. The situation is more nuanced when  $\alpha = 1.5$  and depends on the form of the slowly varying function  $L(\cdot)$ .
  - (b) For  $\alpha \leq 1.5$ , each replication of the estimator terminate in  $O(b)$  time on an average and we require only  $O(1)$  replications, thus resulting in overall complexity of  $O(b)$ . The relative deviation (the ratio of the absolute difference between the estimator and the true value with the true value) of the values returned by the algorithm is well within the specified limits with high probability, even though the estimator variance is infinite.

The above results for infinite increment variance, and in particular the bottleneck arising at  $\alpha = 1.5$ , closely mirror the results proved in Blanchet and Liu (2012) where vastly different state-dependent algorithms are considered.

A brief discussion on practical applications and a literature review may be in order: Efficient estimation of the level crossing probability is important in many practical contexts, e.g., in

computing steady state probability of large delays in  $GI/GI/1$  queues and in ruin probabilities in insurance settings (see, e.g., Asmussen and Glynn (2007)). Siegmund (1976) provides the first weakly efficient importance sampling algorithm for estimating the level crossing probabilities when the increments  $X_n$  are light-tailed using large deviations based exponentially twisted change of measure. Sadowsky and Bucklew (1990) develop a weakly efficient algorithm for estimating  $\mathbb{P}(S_n > na)$  for  $a > 0$ , and  $X_i$  light-tailed, again using exponential twisting based importance sampling distribution (also see Sadowsky (1996), Dupuis and Wang (2004), Blanchet et al. (2009), Dieker and Mandjes (2005) and Agarwal et al. (2013) for related analysis). This problem is important mainly because it forms a building block to many more complex rare event problems involving combination of renewal processes: for examples in queueing, see Parekh and Walrand (1989) and in financial credit risk modeling, see Glasserman and Li (2005) and Bassamboo et al. (2008). Research on efficient simulation of rare events involving heavy-tailed variables first focussed on probabilities such as  $\mathbb{P}\{S_N > b\}$  in the simpler asymptotic regime where  $N$  is fixed or geometrically distributed and  $b \nearrow \infty$ . In this simpler setting state-independent algorithms are easily designed (see, e.g., Asmussen et al. (2000), Juneja and Shahabuddin (2002), Asmussen and Kroese (2006)). In Rajhaa and Juneja (2012), it is shown that a variant capped exponential twisting based state-independent importance sampling, which does not involve any decomposition, provides a strongly efficient estimator for the large deviations probability that we consider in this paper.

Statistical analysis reveals that heavy-tailed distributions are very common in practice: in particular, heavy-tailed increments with infinite variance are a convenient means to explain the long-range dependence observed in tele-traffic data, and to model highly variable claim sizes in insurance settings. Popular references to this strand of literature include Embrechts et al. (1997), Resnick (1997), and Adler et al. (1998).

The organization of the remaining paper is as follows: In Section 2 we discuss preliminary concepts relevant to the problems addressed. We propose our importance sampling method for estimating the large deviations probability and prove its efficiency in Section 3. In Section 4, we develop algorithms for estimating the level crossing probability. Proofs of some of the key results pertaining to efficiency of proposed algorithms and their expected termination time are given in Section 5. Numerical experiments supporting our algorithms are given in Section 6 followed by a brief conclusion in Section 7. Some of the more technical proofs are presented in the appendix.

## 2. PRELIMINARY BACKGROUND

In this section we briefly review the use of importance sampling in estimating rare event probabilities, and the well-known asymptotics for relevant tail probabilities in the existing literature. Throughout this paper, we use Landau's notation for describing asymptotic behaviour of functions: for given functions  $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  and  $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , we say  $f(x) = O(g(x))$  if there exists  $c_1 > 0$  and  $x_1$  large enough such that  $f(x) \leq c_1 g(x)$  for all  $x > x_1$ ; and  $f(x) = \Omega(g(x))$  if there exists  $c_2 > 0$  and  $x_2$  large enough such that  $f(x) \geq c_2 g(x)$  for all  $x > x_2$ . We use  $f(x) = o(g(x))$  if  $f(x)/g(x) \rightarrow 0$ , and  $f(x) \sim g(x)$  if  $f(x)/g(x) \rightarrow 1$ , as  $x \nearrow \infty$ .

**2.1. Rare event simulation and importance sampling.** Let  $A$  denote a rare event on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , i.e.,  $z := \mathbb{P}(A) > 0$  is small (in our setup  $A$  corresponds to the events  $\{S_n > b\}$  or  $\{\tau_b < \infty\}$ ). Suppose that we are interested in obtaining an estimator  $\hat{z}$  for  $z$  such that the relative deviation  $|\hat{z} - z|/z < \epsilon$ , with probability at least  $1 - \delta$ , for given  $\epsilon$  and  $\delta > 0$ . Naive simulation for estimating  $z$  involves drawing  $N$  independent samples of the indicator  $\mathbb{I}_A$  and taking their sample mean as the estimator. For a different measure  $\tilde{\mathbb{P}}(\cdot)$  such

that the Radon-Nikodym derivative  $d\mathbb{P}/d\tilde{\mathbb{P}}$  is well defined on  $A$ , we get:

$$\mathbb{P}(A) = \int_A \frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}(\omega) d\tilde{\mathbb{P}}(\omega) = \tilde{\mathbb{E}}[L\mathbb{I}_A],$$

where  $L := d\mathbb{P}/d\tilde{\mathbb{P}}$  and  $\tilde{\mathbb{E}}[\cdot]$  is the expectation associated with  $\tilde{\mathbb{P}}(\cdot)$ . Define  $Z := L\mathbb{I}_A$ ; then  $Z$  is an unbiased estimator of  $z$  under measure  $\tilde{\mathbb{P}}(\cdot)$ . If  $N$  i.i.d samples  $Z_1, \dots, Z_N$  of  $Z$  are drawn from  $\tilde{\mathbb{P}}(\cdot)$ , then by the strong law of large numbers we have:

$$\hat{z}_N := \frac{Z_1 + \dots + Z_N}{N} \rightarrow z \text{ a.s.},$$

as  $N \nearrow \infty$ . This method of arriving at an estimator is called *importance sampling* (IS). The measure  $\tilde{\mathbb{P}}(\cdot)$  is called the importance sampling measure and  $Z$  is called an importance sampling estimator. Using Chebyshev's inequality allows us to find an upper bound on the number of replications  $N$  required to achieve the desired relative precision:

$$\mathbb{P}\left(\frac{|\hat{z}_N - z|}{z} > \epsilon\right) \leq \frac{\text{Var}(\hat{z}_N)}{z^2 \epsilon^2} = \frac{CV^2(Z)}{N \epsilon^2}.$$

Here  $CV(Z) = \sqrt{\text{Var}(Z)}/z$  is the coefficient of variation of  $Z$ . This enables us to conclude that if we generate at least

$$N = \frac{CV^2(Z)}{\delta \epsilon^2} \tag{1}$$

i.i.d. samples of  $Z$  for computing  $\hat{z}_N$ , we can guarantee the desired relative precision. In naive simulation we use the measure  $\mathbb{P}(\cdot)$  itself and have  $Z = \mathbb{I}_A$  as the estimator; so the number of samples required in (1) grows (roughly proportional to  $z^{-1}$ ) to infinity if  $z \searrow 0$ . As is well known, the choice  $\mathbb{P}^*(\cdot) := \mathbb{P}(\cdot|A)$  as an importance sampling measure yields zero variance for the associated estimator  $Z = z\mathbb{I}_A$  (see e.g., Asmussen and Glynn (2007)). Then, every sample obtained in simulation equals  $z$  with  $\mathbb{P}^*(\cdot)$  probability 1. However, the explicit dependence of  $Z$  on  $z$ , the quantity which we want to estimate, makes this method impractical.

**2.2. Efficiency notions of algorithms.** Consider a family of events  $\{A_n : n \geq 1\}$  such that  $z_n := \mathbb{P}(A_n) \searrow 0$  as the rarity parameter  $n \nearrow \infty$ . For an importance sampling algorithm to compute  $(z_n : n \geq 1)$ , we come up with a sequence of changes of measure  $(\tilde{\mathbb{P}}_n(\cdot) : n \geq 1)$  and estimators  $(Z_n : n \geq 1)$  such that  $\tilde{\mathbb{E}}_n Z_n = z_n$ , where  $\tilde{\mathbb{E}}_n[\cdot]$  denotes the expectation operator under  $\tilde{\mathbb{P}}_n(\cdot)$ .

**Definition 1.** The sequence  $(Z_n : n \geq 1)$  of unbiased importance sampling estimators of  $\{z_n : n \geq 1\}$ , is said to achieve asymptotically vanishing relative error if,

$$\overline{\lim}_{n \rightarrow \infty} \frac{\tilde{\mathbb{E}}_n [Z_n^2]}{z_n^2} \leq 1. \tag{2}$$

The sequence  $(Z_n : n \geq 1)$  is said to be strongly efficient if,

$$\overline{\lim}_{n \rightarrow \infty} \frac{\tilde{\mathbb{E}}_n [Z_n^2]}{z_n^2} < \infty, \tag{3}$$

and weakly efficient if for all  $\epsilon > 0$ ,

$$\overline{\lim}_{n \rightarrow \infty} \frac{\tilde{\mathbb{E}}_n [Z_n^2]}{z_n^{2-\epsilon}} < \infty. \tag{4}$$

The significance of these definitions can be seen from (1): if an algorithm is strongly efficient, the number of simulation runs required to guarantee the desired relative precision stays bounded as  $n \nearrow \infty$ . If  $\text{Var}(Z_n) = o(z_n^2)$ , then  $(Z_n : n \geq 1)$  satisfies asymptotically vanishing relative error property. As a result, it is enough to generate  $o(\delta^{-1}\epsilon^{-2})$  i.i.d. replications of the estimator. As is apparent from the definition, all strongly efficient algorithms are weakly efficient, and vanishing relative error is the strongest notion among all three. Also it can be verified that naive simulation is not even weakly efficient.

**2.3. Related asymptotics.** In this section, we list the well-known asymptotics of the quantities of interest; these asymptotic representations will be useful for arriving at importance sampling measures and proving their efficiency.

1. Recall that  $\beta := (\alpha \wedge 2)^{-1}$ . Then we have,

$$\mathbb{P}\{S_n > b\} \sim n\bar{F}(b), \text{ as } n \nearrow \infty \quad (5)$$

for  $b > n^{\beta+\epsilon}$ ,  $\epsilon > 0$ . (see Mikosch and Nagaev (1998) and references therein). Additionally, the following relations can be found in Mikosch and Nagaev (1998): as  $n \nearrow \infty$ ,

$$\begin{aligned} \mathbb{P}\left\{S_n > b, \max_{k \leq n} X_k < b\right\} &= o(n\bar{F}(b)), \\ \sup_{b > n^{\beta+\epsilon}} |\mathbb{P}\{\#\{1 \leq i \leq n : X_i > b\} = 1 | S_n \geq b\} - 1| &= o(1), \text{ and} \\ \sup_{b > n^{\beta+\epsilon}} \left| \mathbb{P}\left\{\max_{k \leq n-1} X_k \leq b, S_n \geq b | X_n > b\right\} - 1 \right| &= o(1). \end{aligned} \quad (6)$$

These large deviations asymptotics reveal that with the number of summands growing to infinity, with high probability, the sum becomes large because one of the component increments becomes large.

2. Recall that  $\tau_b := \inf\{k : S_k > b + k\mu\}$  and  $M := \sup_n (S_n - n\mu)$ ; the events  $\{M > b\}$  and  $\{\tau_b < \infty\}$  are the same. Let  $\bar{F}_I(\cdot)$  denote the integrated tail of  $\bar{F}(\cdot)$  as below:

$$\bar{F}_I(x) := \int_x^\infty \bar{F}(u) du, \text{ for } x \geq 0.$$

The following asymptotics are well-known (see, for e.g., in Foss et al. (2011) and references therein). As  $b \nearrow \infty$ , uniformly for any positive integer  $n$ ,

$$\begin{aligned} \mathbb{P}\{\tau_b < n\} &\sim \frac{1}{\mu} \int_b^{b+n\mu} \bar{F}(u) du, \text{ and} \\ \mathbb{P}\{\tau_b < \infty\} &\sim \frac{1}{\mu} \bar{F}_I(b). \end{aligned} \quad (7)$$

Then for any positive integers  $n_1$  and  $n_2$  with  $n_1 < n_2$ ,

$$\mathbb{P}\{n_1 < \tau_b \leq n_2\} \sim \frac{1}{\mu} \int_{b+n_1\mu}^{b+n_2\mu} \bar{F}(u) du = \frac{\bar{F}_I(b+n_1\mu) - \bar{F}_I(b+n_2\mu)}{\mu}, \quad (8)$$

uniformly in  $n_1$  and  $n_2$ , as  $b \nearrow \infty$ .

Further, the following characterization of the zero-variance measure  $\mathbb{P}\{\cdot | \tau_b < \infty\}$ , as in Theorem 1.1 of Asmussen and Kluppelberg (1996), sheds light on how the first passage



over a level  $b$  happens asymptotically: If we use  $a(b) := \bar{F}_I(b)/\bar{F}(b)$ , then conditional on  $\tau_b < \infty$ ,

$$\left( \frac{\tau_b}{a(b)}, \left( \frac{S_{\lfloor u\tau_b \rfloor}}{\tau_b} : 0 \leq u < 1 \right), \frac{S_{\tau_b} - b}{a(b)} \right) \Rightarrow \left( \frac{Y_0}{\mu}, (-u\mu : 0 \leq u < 1), Y_1 \right) \quad (9)$$

in  $\mathbb{R} \times D[0, 1) \times \mathbb{R}$ . The joint law of  $Y_0, Y_1$  is defined as follows: for  $y_0, y_1 \geq 0$ ,  $\mathbb{P}\{Y_0 > y_0, Y_1 > y_1\} = \mathbb{P}\{Y_1 > y_0 + y_1\}$  with  $Y_0 \stackrel{d}{=} Y_1$ , and

$$\mathbb{P}\{Y_1 > y_1\} = \frac{1}{(1 + y_1/(\alpha - 1))^{\alpha-1}}.$$

Now we state a part of Karamata's theorem that provides an asymptotic characterization of the integrated tails of regularly varying functions: Consider a regularly varying function  $V(\cdot)$  with index  $-\alpha$ ; if  $\beta$  is such that  $\alpha - \beta > 1$ ,

$$\int_x^\infty u^\beta V(u) du \sim \frac{x^{\beta+1} V(x)}{\alpha - \beta - 1}, \text{ as } x \nearrow \infty. \quad (10)$$

See Embrechts et al. (1997) or Borovkov and Borovkov (2008) for further details.

### 3. SIMULATION OF $\{S_n > b\}$

Let  $X$  be a zero mean random variable with distribution  $F(\cdot)$  satisfying the following:

**Assumption 1.** *The tail probabilities are given by  $\bar{F}(x) := \mathbb{P}\{X > x\} = x^{-\alpha} L(x)$ , for some slowly varying function  $L(\cdot)$  and  $\alpha > 1$ . If  $\text{Var}[X] = \infty$ , then*

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}\{X < -x\}}{\mathbb{P}\{X > x\}} < \infty.$$

For the independent collection  $(X_n : n \geq 1)$  of random variables which are distributed identically as  $X$ , define the random walk  $(S_n : n \geq 0)$  as below:

$$S_0 = 0, \text{ and } S_n = X + 1 + \dots + X_n \text{ for } n \geq 1.$$

In this section we devise a simulation procedure for estimating the large deviation probabilities  $\mathbb{P}\{S_n > b\}$  for  $b > n^{\beta+\epsilon}$  given any  $\epsilon > 0$ , and prove its efficiency as  $n \nearrow \infty$ . Recall that  $\beta := (\alpha \wedge 2)^{-1}$ . The strategy is to partition the event  $\{S_n > b\}$  based on whether the maximum of the increments  $\{X_1, \dots, X_n\}$  has exceeded the large value  $b$  or not (Juneja 2007 considers this approach when  $n$  is fixed) :

$$A_{\text{dom}}(n, b) := \left\{ S_n > b, \max_{k \leq n} X_k \geq b \right\} \text{ and } A_{\text{res}}(n, b) := \left\{ S_n > b, \max_{k \leq n} X_k < b \right\}.$$

The asymptotics (5) and (6) in Section 2.3 indicate that for large values of  $n$ , the most likely way for the sum  $S_n$  to exceed  $b$  is to have at least one of the increments  $\{X_1, \dots, X_n\}$  exceed  $b$ . Hence the probability of the event  $A_{\text{res}}$  is vanishingly small compared to the probability of  $A_{\text{dom}}$ , as  $n \nearrow \infty$ ; the suffixes stand to indicate that  $A_{\text{dom}}$  is the dominant way of occurrence of  $\{S_n > b\}$  for large  $n$ , and the other event has only residual contributions. We estimate  $\mathbb{P}(A_{\text{dom}})$  and  $\mathbb{P}(A_{\text{res}})$  independently via different changes of measure that typify the way in which the respective events occur, and add the individual estimates to arrive at the final estimator for  $\mathbb{P}\{S_n > b\}$ .

**3.1. Simulating  $A_{\text{dom}}$ .** For the simulation of  $A_{\text{dom}}$ , we follow the two-step procedure outlined in Chan et al. (2012):

1. Choose an index  $I$  uniformly at random from  $\{1, \dots, n\}$
2. For  $k = 1, \dots, n$ , generate a realization of  $X_k$  from  $F(\cdot | X_k \geq b)$  if  $k = I$ ; otherwise, generate  $X_k$  from  $F(\cdot)$ .

Let  $\tilde{\mathbb{P}}(\cdot)$  denote the measure induced when the increments are generated according to the above procedure, and let  $\tilde{\mathbb{E}}[\cdot]$  denote the corresponding expectation operator. Note that the probability measure  $\mathbb{P}(\cdot)$  is absolutely continuous with respect to  $\tilde{\mathbb{P}}(\cdot)$  when restricted to  $A_{\text{dom}}$ . We have,

$$d\tilde{\mathbb{P}}(x_1, \dots, x_n) = \sum_{k=1}^n \frac{1}{n} \cdot \frac{dF(x_1) \dots dF(x_n)}{\bar{F}(b)} \mathbf{1}(x_k \geq b).$$

Therefore the likelihood ratio on the set  $A_{\text{dom}}$  is given by,

$$\frac{d\mathbb{P}}{d\tilde{\mathbb{P}}}(X_1, \dots, X_n) = \frac{n\bar{F}(b)}{\#\{X_i \geq b : 1 \leq i \leq n\}},$$

and the resulting unbiased estimator for the evaluation of  $\mathbb{P}(A_{\text{dom}})$  is,

$$Z_{\text{dom}}(n, b) := \frac{n\bar{F}(b)}{\#\{X_i \geq b : 1 \leq i \leq n\}} \mathbb{I}(A_{\text{dom}}). \quad (11)$$

Generate  $N$  independent realizations of  $Z_{\text{dom}}$  and take their sample mean as an estimator of  $\mathbb{P}(A_{\text{dom}})$ . To evaluate how large  $N$  should be chosen so that the computed estimate satisfies the given relative error specification, we need to obtain bounds on the variance of  $Z_{\text{dom}}$ . Since  $\#\{X_i \geq b : 1 \leq i \leq n\}$  is at least 1, when the increments are drawn following the measure  $\tilde{\mathbb{P}}(\cdot)$ , we have:  $Z_{\text{dom}}(n, b) \leq n\bar{F}(b)$ , and hence,

$$\tilde{\mathbb{E}}[Z_{\text{dom}}^2(n, b)] \leq (n\bar{F}(b))^2.$$

Also  $\tilde{\mathbb{E}}[Z_{\text{dom}}(n, b)] = \mathbb{P}(A_{\text{dom}}(n, b)) \sim \mathbb{P}\{S_n > b\} \sim n\bar{F}(b)$ , as  $n \nearrow \infty$ . Therefore we get,

$$\text{Var}[Z_{\text{dom}}(n, b)] = o\left((n\bar{F}(b))^2\right), \text{ as } n \nearrow \infty. \quad (12)$$

**3.2. Simulating  $A_{\text{res}}$ .** We see that all the increments  $\{X_1, \dots, X_n\}$  are bounded from above by  $b$  on the occurrence of event  $A_{\text{res}}$ . Though the bound on the increments vary with  $n$ , we can employ methods similar to exponential twisting of light-tailed random walks to simulate the event  $A_{\text{res}}$ , as illustrated in this section. For given  $b$ , define

$$\Lambda_b(\theta) := \log \left( \int_{-\infty}^b \exp(\theta x) F(dx) \right), \quad \theta \geq 0.$$

Since the upper limit of integration is  $b$ ,  $\Lambda(\cdot)$  is well-defined for any positive value of  $\theta$ . For given values of  $n$  and  $b$ , consider the distribution function  $F_\theta(\cdot)$  satisfying,

$$\frac{dF_\theta(x)}{dF(x)} = \exp(\theta_n x - \Lambda_b(\theta_n, b)) \mathbf{1}(x < b),$$

for all  $x \in \mathbb{R}$  and some  $\theta_{n,b} > 0$ . Now the prescribed procedure is to just obtain independent samples of the increments  $\{X_1, \dots, X_n\}$  from  $F_\theta(\cdot)$  and compute the likelihood ratio due to the



procedure of sampling from a different distribution  $F_\theta(\cdot)$ . Let  $\mathbb{P}_\theta(\cdot)$  and  $\mathbb{E}_\theta[\cdot]$  denote, respectively, the corresponding importance sampling change of measure and its associated expectation operator. Note that the dependence of  $F_\theta(\cdot)$ ,  $\mathbb{P}_\theta(\cdot)$  and  $\mathbb{E}_\theta[\cdot]$  on  $n$  and  $b$  has been suppressed in the notation. Then for given values of  $n$  and  $b$ , we have the following unbiased estimator for the computation of  $\mathbb{P}(A_{\text{res}})$  :

$$Z_{\text{res}}(n, b) := \exp(-\theta_n S_n + n\Lambda_b(\theta_{n,b})) \mathbb{I}(A_{\text{res}}). \quad (13)$$

Now generate independent replications of  $Z_{\text{res}}$  and take their sample mean as the computed estimate for  $\mathbb{P}(A_{\text{res}})$ . However it remains to choose  $\theta_{n,b}$ . Since  $S_n$  is larger than  $b$  on  $A_{\text{res}}$ ,

$$Z_{\text{res}}(n, b) \leq \exp(-\theta_n b + n\Lambda_b(\theta_{n,b})) \mathbb{I}(A_{\text{res}}).$$

If we choose

$$\theta_{n,b} := -\frac{\log(n\bar{F}(b))}{b}, \text{ then} \quad (14)$$

$$Z_{\text{res}}(n, b) \leq n\bar{F}(b) \exp(n\Lambda_b(\theta_{n,b})) \mathbb{I}(A_{\text{res}}). \quad (15)$$

We use Lemma 1, which is proved in the appendix, to obtain an upper bound on the second moment of the estimator  $Z_{\text{res}}$ .

**Lemma 1.** *For the choice of  $\theta_{n,b}$  as in (14),*

$$\exp(\Lambda_b(\theta_{n,b})) \leq 1 + \frac{1}{n}(1 + o(1)),$$

as  $n \nearrow \infty$ , uniformly for  $b > n^{\beta+\epsilon}$ .

Therefore there exists a constant  $c$  such that

$$\exp(n\Lambda_b(\theta_{n,b})) \leq c,$$

for all admissible values of  $n$  and  $b$ . We evaluate the second moment of the estimator  $Z_{\text{res}}$  through the equivalent expectation operation corresponding to the original measure  $\mathbb{P}(\cdot)$  as below:

$$\mathbb{E}_\theta [Z_{\text{res}}^2(n, b)] = \mathbb{E} [Z_{\text{res}}(n, b)] \leq cn\bar{F}(b)\mathbb{P}(A_{\text{res}}),$$

where the last inequality follows from (15). Since  $\mathbb{P}(A_{\text{res}}) = o(n\bar{F}(b))$ , as in (6), we obtain that:

$$\text{Var} [Z_{\text{res}}(n, b)] = o((n\bar{F}(b))^2), \text{ as } n \nearrow \infty, \quad (16)$$

thus arriving at the following theorem:

**Theorem 1.** *If the realizations of the estimators  $Z_{\text{dom}}$  and  $Z_{\text{res}}$  are generated respectively from the measures  $\tilde{\mathbb{P}}(\cdot)$  and  $\mathbb{P}_\theta(\cdot)$ , and if we let,*

$$Z(n, b) := Z_{\text{dom}}(n, b) + Z_{\text{res}}(n, b),$$

*then under Assumption 1, the family of estimators  $(Z(n, b) : n \geq 1, b > n^{\beta+\epsilon})$  achieves asymptotically vanishing relative error for the estimation of  $\mathbb{P}\{S_n > b\}$ , as  $n \nearrow \infty$ ; that is,*

$$\frac{\text{Var} [Z(n, b)]}{(\mathbb{P}\{S_n > b\})^2} = o(1),$$

as  $n \nearrow \infty$ , uniformly for  $b > n^{\beta+\epsilon}$ .

*Proof.* Since the realizations of  $Z_{\text{dom}}$  and  $Z_{\text{res}}$  are generated independent of each other, the variance of  $Z$  is just the sum of variances of  $Z_{\text{dom}}$  and  $Z_{\text{res}}$ ; the proof is now evident from (12), (16) and (5).  $\square$

*Remark 1.* A consequence of the above theorem is that, due to (1), the number of i.i.d. replications of  $Z(n, b)$  required to achieve  $\epsilon$ -relative precision with probability at least  $1 - \delta$  is at most  $o(\epsilon^{-2}\delta^{-1})$ , independent of the rarity parameters  $n$  and  $b$ . In our algorithm each replication demands  $O(n)$  computational effort, thus requiring a overall computational cost of  $O(n)$ , as  $n \nearrow \infty$ .

*Remark 2.* One can easily check that, this same simulation procedure can also be used to efficient compute probabilities  $P(S_N > b)$  when  $N$  is a random variable independent of  $\{X_i\}$ .

#### 4. SIMULATION METHODOLOGY FOR $\{\tau_b < \infty\}$

As before, the sequence  $(S_n : n \geq 0)$  with  $S_0 := 0$  and  $S_n := X_1 + \dots + X_n$  represents the random walk associated with the i.i.d collection  $(X_n : n \geq 1)$ . We have  $\mathbb{E}X_n = 0$ , and  $\mathbb{P}\{X_n > x\} = x^{-\alpha}L(x)$  for some slowly varying function  $L(\cdot)$  and  $\alpha > 1$ . Given  $\mu > 0$ ,  $M := \sup_n (S_n - n\mu)$ . Since  $(S_n - n\mu : n \geq 0)$  is a random walk with negative drift, the random variable  $M$  is proper. For  $b > 0$ , recall that the first-passage time  $\tau_b := \inf\{n \geq 0 : S_n - n\mu > b\}$ . In this section we present simulation methods for the efficient computation of

$$\mathbb{P}\{M > b\} = \mathbb{P}\{\tau_b < \infty\}, \text{ as } b \nearrow \infty.$$

Naive simulation of  $\{\tau_b < \infty\}$  will require generation of all the increments until the maximum of the partial sums exceed  $b$ . Due to the negative drift of the random walk  $(S_n - n\mu : n \geq 0)$ , we have  $\tau_b \nearrow \infty$  a.s. as  $b \nearrow \infty$ , and hence this method is not computationally feasible. To counter the prospect of generating uncontrollably large number of increment random variables in simulation, we re-express  $\mathbb{P}\{\tau_b < \infty\}$  as below: Consider a strictly increasing sequence of integers  $(n_k : k \geq 0)$  with  $n_0 = 0$ ; also fix  $p := (p_k : k \geq 1)$  satisfying  $p_k > 0$  for all  $k$  and  $\sum_k p_k = 1$ ; the vector  $p$  can be seen as a probability mass function on positive integers. If we consider an auxiliary random variable  $K$  which takes the value of positive integer  $k$  with probability  $p_k$ , then we can write,

$$\begin{aligned} \mathbb{P}\{\tau_b < \infty\} &= \sum_{k \geq 1} p_k \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}}{p_k} \\ &= \mathbb{E} \left[ \frac{\mathbb{P}\{n_{K-1} < \tau_b \leq n_K\}}{p_K} \right], \end{aligned} \tag{17}$$

where  $\mathbb{E}[f(K)] = \sum_{k \geq 1} p_k f(k)$ , for any  $f : \mathbb{Z}^+ \rightarrow \mathbb{R}$ .

Now in a simulation run, if the realized value of the auxiliary random variable  $K$  is  $k$ , generate a sample from a probability measure, possibly different from  $\mathbb{P}(\cdot)$ , of a random variable  $Z_k$  that has  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}$  as its expectation under the changed measure. Then equation (17) assures us that repeated simulation runs involving generation of  $K$  according to the measure induced by  $p$ , and taking sample mean of such realizations of  $Z_K/p_K$  following the changed measure will yield an unbiased estimator for the quantity  $\mathbb{P}\{\tau_b < \infty\}$ .

The performance of any importance sampling algorithm following the outlined procedure will depend crucially on the choice of probabilities  $p_k$ , and the change of measure employed to estimate  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}$ , for  $k \geq 1$ . The sequence  $(n_k : k \geq 0)$  partitions non-negative

integers into ‘blocks’  $((n_{k-1}, n_k] : k \geq 1)$ . For reasons that will be clear later, we choose the blocks  $(n_{k-1}, n_k]$  in the following manner: Fix a positive integer  $r > 1$  and let,

$$n_0 = 0, n_k = r^k, \text{ for } k \geq 1.$$

In the following section, we detail the importance sampling schemes for the efficient computation of the quantities  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}, k \geq 1$ ; these will be used as building blocks to efficiently compute the ultimate object of interest  $\mathbb{P}\{\tau_b < \infty\}$ .

**4.1. Efficient simulation of  $\{n_{k-1} < \tau_b \leq n_k\}$ .** In this section we present our state-independent importance sampling procedure, for the efficient computation of the probabilities  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}$ , that are uniformly efficient for  $k \geq 1$ . Define the following events:

$$A_k = \bigcup_{i=n_{k-1}+1}^{n_k} \{X_i > b + i\mu\} \text{ and } B_k = \bigcap_{i=1}^{n_k} \{X_i < b + n_{k-1}\mu\}.$$

The events  $A_k$  and  $B_k$  are defined in the same spirit as that of  $A_{\text{dom}}$  and  $A_{\text{res}}$  in the simulation of  $\{S_n > b\}$  in Section 3: the event  $A_k$  includes sample paths that have at least one ‘‘big’’ jump of appropriate size in one of the increments indexed between  $n_{k-1}$  and  $n_k$ , whereas on the other set  $B_k$  we have all the increments bounded from above. The following lemma, proved in the appendix, asserts that asymptotically  $A_k$  is the most likely way for the event  $\{n_{k-1} < \tau_b \leq n_k\}$  to happen.

**Lemma 2.** *For any  $\epsilon > 0$ , there exists  $b_\epsilon$  such that for all  $b > b_\epsilon$ ,*

$$\sup_{k \geq 1} \left| \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\}}{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}} - 1 \right| < \epsilon.$$

As in the simulation of large deviation probabilities of sums of random variables in Section 3, we can partition the event  $\{n_{k-1} < \tau_b \leq n_k\}$  into:

$$\{n_{k-1} < \tau_b \leq n_k, A_k\}, \{n_{k-1} < \tau_b \leq n_k, B_k\} \text{ and } \{n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k\},$$

and arrive at unbiased estimators for their probabilities separately via different importance sampling measures.

**4.1.1. Simulating  $\{n_{k-1} < \tau_b \leq n_k, A_k\}$ .** Let  $q_k(b) := \sum_{i=n_{k-1}+1}^{n_k} \bar{F}(b + i\mu)$ . We prescribe the following two step procedure:

1. Choose an index  $J \in \{n_{k-1} + 1, \dots, n_k\}$  such that  $\Pr\{J = n\} = \frac{\bar{F}(b+n\mu)}{q_k(b)}$ , for  $n_{k-1} < n \leq n_k$ .
2. Simulate the increment  $X_n$  from  $F(\cdot | X_n \geq b + n\mu)$ , if  $n = J$ ; otherwise, simulate  $X_n$  from  $F(\cdot)$ , for any  $n \leq n_k$ .

In this sampling procedure, we induce the ‘big’ jumps typically responsible for the occurrence of  $\{n_{k-1} < \tau_b \leq n_k\}$  with suitable probabilities by sampling from the conditional distribution  $F(\cdot | X_J \geq b + J\mu)$ . This sampling procedure results in the importance sampling measure  $\mathbb{P}_{k,1}(\cdot)$  characterised by:

$$d\mathbb{P}_{k,1}(x_1, \dots, x_{n_k}) := \sum_{i=n_{k-1}+1}^{n_k} \frac{\bar{F}(b + i\mu)}{q_k(b)} \cdot \frac{dF(x_1) \dots dF(x_{n_k})}{\bar{F}(b + i\mu)} \mathbf{1}(x_i \geq b + i\mu).$$

This in turn yields a likelihood ratio,

$$\frac{d\mathbb{P}}{d\mathbb{P}_{k,1}}(X_1, \dots, X_{n_k}) = \frac{q_k(b)}{\#\{X_i \geq b + i\mu : n_{k-1} < i \leq n_k\}},$$

on the set  $A_k$ . Then we have,

$$Z_{k,1}(b) := \frac{q_k(b)}{\#\{X_i \geq b + i\mu : n_{k-1} < i \leq n_k\}} \mathbb{I}(n_{k-1} < \tau_b \leq n_k, A_k) \quad (18)$$

as the unbiased estimator for the quantity  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\}$ . Here note that  $\mathbb{I}(n_{k-1} < \tau_b \leq n_k, A_k) = 1$  a.s. under  $\mathbb{P}_{k,1}$ .

**Lemma 3.** *Uniformly for  $k \geq 1$ ,*

$$\text{Var}[Z_{k,1}(b)] = o\left((\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2\right), \text{ as } b \nearrow \infty.$$

*Proof.* Since the quantity  $\#\{X_i \geq b + i\mu : n_{k-1} < i \leq n_k\}$  is at least 1 when the increments are generated from  $\mathbb{P}_{k,1}(\cdot)$ ,

$$Z_{k,1}(b) \leq q_k(b),$$

and hence,

$$\mathbb{E}_{k,1}[Z_{k,1}^2] \leq q_k^2(b). \quad (19)$$

We have,

$$q_k(b) = \sum_{i=n_{k-1}+1}^{n_k} \bar{F}(b + i\mu) \leq \sum_{i=n_{k-1}+1}^{n_k} \int_{i-1}^i \bar{F}(b + u\mu) du = \int_{n_{k-1}}^{n_k} \bar{F}(b + u\mu) du.$$

Changing variables from  $u$  to  $v = b + u\mu$  gives,

$$q_k(b) \leq \frac{1}{\mu} \int_{b+n_{k-1}\mu}^{b+n_k\mu} \bar{F}(v) dv.$$

Now given  $\epsilon > 0$ , because of (8) and (19),

$$\mathbb{E}_{k,1}[Z_{k,1}^2] \leq (1 + \epsilon) (\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2$$

for all  $k$  and  $b$  large enough. Also,

$$\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\} \geq (1 - \epsilon) \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\},$$

for all  $k$ , because of Lemma 2. Therefore,

$$\begin{aligned} \text{Var}[Z_{k,1}(b)] &\leq (1 + \epsilon - (1 - \epsilon)^2) (\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2 \\ &\leq 3\epsilon (\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2. \end{aligned} \quad \square$$

**4.1.2. Simulating**  $\{n_{k-1} < \tau_b \leq n_k, B_k\}$ . On the event  $B_k$ , none of the random variables  $X_1, \dots, X_{n_k}$  exceed the level  $(b + n_{k-1}\mu)$ ; since these increments are bounded (on  $B_k$ ), we can draw their samples from an appropriately truncated, exponentially twisted variation of  $F(\cdot)$  without losing absolute continuity on  $\{n_{k-1} < \tau_b \leq n_k, B_k\}$ . For estimating  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, B_k\}$ , we draw samples of  $X_1, \dots, X_{\tau_b \wedge n_k}$  independently from the distribution  $F_k(\cdot)$  satisfying,

$$\frac{dF_k(x)}{dF(x)} = \exp(\theta_k x - \Lambda_k(\theta_k)) \mathbf{1}(x < b + n_{k-1}\mu), \quad x \in \mathbb{R};$$

$$\text{here, } \theta_k(= \theta_k(b)) := \frac{-\log(n_k \bar{F}(b + n_{k-1}\mu))}{b + n_{k-1}\mu}, \text{ and} \quad (20)$$

$$\Lambda_k(\theta) := \log \left( \int_{-\infty}^{b+n_{k-1}\mu} \exp(\theta_k x) F(dx) \right), \quad \theta \geq 0. \quad (21)$$

Let  $\mathbb{P}_{k,2}(\cdot)$  be the measure induced by drawing samples as above. Then the resulting likelihood ratio on  $\{n_{k-1} < \tau_b \leq n_k, B_k\}$  is:

$$\frac{d\mathbb{P}}{d\mathbb{P}_{k,2}}(X_1, \dots, X_{n_k}) = \exp(-\theta_k S_{\tau_b} + \tau_b \Lambda_k(\theta_k)).$$

The associated estimator for computing  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, B_k\}$  is:

$$Z_{k,2}(b) := \exp(-\theta_k S_{\tau_b} + \tau_b \Lambda_k(\theta_k)) \mathbb{I}(n_{k-1} < \tau_b \leq n_k, B_k) \quad (22)$$

The following uniform bounds, which help in analyzing the variance of the estimator  $Z_{k,2}$ , are proved in the appendix.

**Lemma 4.** *For all values of  $k$  and  $b$ , there exists a positive constant  $c_1$  such that,*

$$\exp(n_k \Lambda_k(\theta_k)) \leq c_1.$$

**Lemma 5.** *For all values of  $k$  and  $b$ , there exists a positive constant  $c_2$  such that,*

$$\frac{n_k \bar{F}(b + n_{k-1}\mu)}{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}} \leq c_2.$$

Using these results, we now present an asymptotic analysis on the variance of the estimators  $Z_{k,2}(\cdot)$ .

**Lemma 6.** *Uniformly for  $k \geq 1$ ,*

$$\text{Var}[Z_{k,2}(b)] = o\left((\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2\right), \text{ as } b \nearrow \infty.$$

*Proof.* Since  $n_k/r < \tau_b \leq n_k$  on the event  $\{n_{k-1} < \tau_b \leq n_k\}$ ,

$$\exp(\tau_b \Lambda_k(\theta_k)) \mathbb{I}(n_{k-1} < \tau_b \leq n_k, B_k) \leq c_1 \vee c_1^{r^{-1}} =: c$$

**Resolve**  $c_1, c$ . because of Lemma 4. Further note that  $\theta_k S_{\tau_b} \geq -\log(n_k \bar{F}(b + n_{k-1}\mu))$  on  $\{n_{k-1} < \tau_b \leq n_k\}$ . Therefore from (22),

$$Z_{k,2}(b) \leq c (n_k \bar{F}(b + n_{k-1}\mu)) \mathbb{I}(n_{k-1} < \tau_b \leq n_k, B_k), \text{ for all } k.$$

Now, changing the expectation operator in the evaluation of second moment of the estimator, results in the following bound:

$$\begin{aligned}\mathbb{E}_{k,2} [Z_{k,2}^2(b)] &= \mathbb{E} [Z_{k,2}(b)] \\ &\leq c (n_k \bar{F}(b + n_{k-1}\mu)) \mathbb{P}\{n_{k-1} < \tau_b \leq n_k, B_k\}.\end{aligned}$$

Here we apply Lemma 2 for a bound on the probability term in the above expression. Given  $\epsilon > 0$ , for all  $b$  large enough, we have:

$$\mathbb{E}_{k,2} [Z_{k,2}^2(b)] \leq c (n_k \bar{F}(b + n_{k-1}\mu)) (\epsilon \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}), \text{ for all } k.$$

Now using Lemma 5 we obtain,

$$\text{Var} [Z_{k,2}(b)] \leq \epsilon c.c_2 (\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2. \quad \square$$

**4.1.3. Simulating  $\{n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k\}$ .** We draw samples in a two step procedure similar to that in the Section 4.1.1.

1. Choose an index  $J$  uniformly at random from  $\{1, \dots, n_k\}$
2. Simulate the increment  $X_n$  from  $F(\cdot | X_n \geq b + n_{k-1}\mu)$ , if  $n = J$ ; otherwise, simulate  $X_n$  from  $F(\cdot)$ , for any  $n \leq n_k$ .

If  $\mathbb{P}_{k,3}$  denotes the change of measure induced by drawing samples according to the above procedure, then the likelihood ratio on the set  $\{n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k\}$  is:

$$\frac{d\mathbb{P}}{d\mathbb{P}_{k,3}}(X_1, \dots, X_{n_k}) = \frac{n_k \bar{F}(b + n_{k-1}\mu)}{\#\{X_i \geq b + n_{k-1}\mu : 1 < i \leq n_k\}}.$$

The resulting estimator for the computation of  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k\}$  is:

$$Z_{k,3}(b) := \frac{n_k \bar{F}(b + n_{k-1}\mu)}{\#\{X_i \geq b + n_{k-1}\mu : 1 < i \leq n_k\}} \mathbb{I}(n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k). \quad (23)$$

Similar to Lemmas 3 and 6, we have the following result on the variance of  $Z_{k,3}(\cdot)$ :

**Lemma 7.** *Uniformly for  $k \geq 1$ ,*

$$\text{Var} [Z_{k,3}(b)] = o\left((\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2\right), \text{ as } b \nearrow \infty.$$

*Proof.* When the increments are generated as prescribed in the above two-step procedure, we have  $\#\{X_i \geq b + n_{k-1}\mu : 1 < i \leq n_k\} \geq 1$ , and hence,

$$Z_{k,3}(b) \leq n_k \bar{F}(b + n_{k-1}\mu) \mathbb{I}(n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k).$$

Now a bound on the second moment of the estimator can be obtained as before:

$$\begin{aligned}\mathbb{E}_{k,3} [Z_{k,3}^2(b)] &= \mathbb{E} [Z_{k,3}(b)] \\ &\leq n_k \bar{F}(b + n_{k-1}\mu) \mathbb{P}\{n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k\}.\end{aligned}$$

Given  $\epsilon > 0$ , due to application of Lemma 2, for all  $k \geq 1$  and  $b$  large enough, we have:

$$\mathbb{E}_{k,3} [Z_{k,3}^2(b)] \leq n_k \bar{F}(b + n_{k-1}\mu) (\epsilon \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}).$$

Using Lemma 5, we write,

$$\text{Var} [Z_{k,3}(b)] \leq \epsilon c_2 (\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2,$$

thus establishing the claim.  $\square$



The estimator for  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}$  can be obtained by summing the estimators of  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\}$ ,  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, B_k\}$ , and  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k\}$ :

$$Z_k(b) := Z_{k,1}(b) + Z_{k,2}(b) + Z_{k,3}(b).$$

Since the random variables  $\{Z_{k,j}(b), j = 1, 2, 3\}$  are generated independent of each other,

$$\begin{aligned} \text{Var}[Z_k(b)] &= \text{Var}[Z_{k,1}(b)] + \text{Var}[Z_{k,2}(b)] + \text{Var}[Z_{k,3}(b)] \\ &= o\left((\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2\right), \end{aligned}$$

uniformly for  $k \geq 1$ , as  $b \nearrow \infty$ , because of Lemmas 3, 6, and 7. This yields the following theorem:

**Theorem 2.** *The family of estimators  $\{Z_k(b); k \geq 1, b > 0\}$  achieves asymptotically vanishing relative error for the unbiased estimation of  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}$ , uniformly in  $k$ , as  $b \nearrow \infty$ ; that is,*

$$\sup_{k \geq 1} \frac{\text{Var}[Z_k(b)]}{(\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\})^2} = o(1),$$

as  $b \nearrow \infty$ .

*Remark 3.* For our choice of importance sampling measures, the likelihood ratios resulting in the simulation of  $\{n_{k-1} < \tau_b \leq n_k, B_k\}$  and  $\{n_{k-1} < \tau_b \leq n_k, \bar{A}_k \cap \bar{B}_k\}$  are  $O(n_k \bar{F}(b + n_{k-1}\mu))$ . To have vanishing relative error, we need  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}$  to be of the same order, which happens when the choice of  $(n_k : k \geq 0)$  is geometric, as shown in Lemma 5.

**4.2. Simulation of  $\{\tau_b < \infty\}$  - the finite variance case.** Here we develop on the ideas stated at the beginning of Section 4. We have the increasing sequence of integers  $(n_k : k \geq 0)$ ,

$$n_0 = 0, n_k = r^k \text{ for } k \geq 1,$$

for some integer  $r > 1$ . Further, we have an auxiliary random variable  $K$  taking values in positive integers according to the probability mass function  $(p_k : k \geq 1)$ . As in (17), we re-express the quantity of interest as:

$$\mathbb{P}\{\tau_b < \infty\} = \mathbb{E} \left[ \frac{\mathbb{P}\{n_{K-1} < \tau_b \leq n_K\}}{p_K} \right].$$

From Section 4.1, we have estimators  $\{Z_k(b) : k \geq 1\}$  that can be used to compute the corresponding probabilities  $\{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\} : k \geq 1\}$  in an efficient manner. Consider the following simulation procedure:

1. Draw a sample of  $K$  such that  $\Pr\{K = k\} = p_k$ .
2. Generate a realization of  $Z_K(b)$  as in Section 4.1.
3. Return  $\frac{Z_K(b)}{p_K}$ .

We present the sample mean of the values returned by  $N$  independent simulation runs of the above procedure as our final estimate of  $\mathbb{P}\{\tau_b < \infty\}$ . Let  $Q(\cdot)$  denote the probability measure in the path space induced by the generation of increment random variables as a result of one run of this sampling procedure; let  $\mathbb{E}^Q[\cdot]$  be the expectation operator associated with  $Q(\cdot)$ . Given  $b > 0$ , the overall unbiased estimator for the computation of  $\mathbb{P}\{\tau_b < \infty\}$  is,

$$Z(b) := \frac{Z_K(b)}{p_K}.$$

Note that the number of independent simulation runs needed to achieve a desired relative precision, as in (1), is directly related to the sampling variance of  $Z(b)$ . If  $(Z(b) : b > 0)$  offer asymptotically vanishing relative error, we just need  $o(\epsilon^{-2}\delta^{-1})$  independent replications of the estimator. However, as pointed in Hammersley and Handscomb (1965), and further justified in Glynn and Whitt (1992), both the variance of an estimator and the expected computational effort required to generate a single sample are important performance measures, and their product can be considered as a ‘figure of merit’ in comparing performance of algorithms that provide unbiased estimators of  $\mathbb{P}\{\tau_b < \infty\}$ . For any given  $b$ , let  $\nu_b$  denote the largest index of the increment random variables ( $X_i$ s) considered for simulation in a particular simulation run. The expectation of  $\nu_b$ , then gives a measure of the expected number of increment random variables generated, and subsequently of the expected computational effort in every simulation run. In particular, the latter may be bounded from above by a constant  $C > 0$  times the expectation of  $\nu_b$ .

In a single run of the above procedure, if the realized value of  $K$  is  $k$ , we look for estimating  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}$  which does not entail the generation of more than  $n_k$  increment random variables, thus ensuring termination. In particular,  $n_{K-1} \leq \nu_b \leq n_K$ . The following theorems give a measure of both the variance and the expected computational effort per replicaton of  $Z(b)$  for a specific choice of the probabilities  $p_k$ :

**Theorem 3.** *For*

$$p_k = \frac{\bar{F}_I(b + n_{k-1}\mu) - \bar{F}_I(b + n_k\mu)}{\bar{F}_I(b)}, k \geq 1, \quad (24)$$

*the family of unbiased estimators  $(Z(b) : b > 0)$  achieves asymptotically vanishing relative error for the computation of  $\mathbb{P}\{\tau_b < \infty\}$ , as  $b \nearrow \infty$ ; that is:*

$$\overline{\lim}_{b \rightarrow \infty} \frac{\text{Var}^Q[Z(b)]}{\mathbb{P}\{\tau_b < \infty\}^2} = 0.$$

**Theorem 4.** *If  $\bar{F}(\cdot)$  is regularly varying with index  $\alpha > 2$ , for the choice of  $p = (p_k : k \geq 1)$  in (24):*

$$E^Q[\nu_b] \leq \frac{r + o(1)}{\mu(\alpha - 2)}b, \text{ as } b \nearrow \infty.$$

Proofs of both these results are given later in Section 5.

*Remark 4.* From Theorem 3, we have the vanishing relative error property for computing  $\mathbb{P}\{\tau_b < \infty\}$  whenever the increment random variables  $X_n$  have finite mean (irrespective of the variance). Therefore we require only  $o(\epsilon^{-2}\delta^{-1})$  i.i.d replications of  $Z(b)$  to arrive at estimators that differ relatively at most by  $\epsilon$  with probability at least  $1 - \delta$ . Now from Theorem 4 we conclude that, if the tail index  $\alpha > 2$  (in which case the increments have finite variance), our importance sampling methodology estimates  $\mathbb{P}\{\tau_b < \infty\}$  in  $O(b)$  expected computational effort.

*Remark 5.* From the conditional limit result in (9), one can infer that the values  $p_k$  as in (24) match the zero-variance probability  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k | \tau_b < \infty\}$  asymptotically. For tails  $\bar{F}(\cdot)$  with regularly varying index  $1 < \alpha < 2$ , we have that  $\mathbb{E}[\tau_b | \tau_b < \infty] = \infty$ ; that is, the zero-variance measure itself has infinite expected termination time! Since  $p_k$  are assigned a value similar to  $\mathbb{P}\{n_{k-1} < \tau_b \leq n_k | \tau_b < \infty\}$ , one might suspect infinite expected termination time for a single run of Algorithm 1 as well. As we note later in Remark 9 after proof of Theorem 4, for  $p_k$ s as in (24), this is indeed the case.

**4.3. Simulation of  $\{\tau_b < \infty\}$  - the infinite variance case.** As in Remark 5, infinite termination time for a simulation algorithm is clearly unacceptable. The following question then is natural: By choosing  $p_k$ s differently, even if it means compromising on estimator variance, can one achieve finite expected termination time for the procedure in Section 4.2? Before answering this question below, we introduce a family of tail distributions and their integrated counterparts: for any  $\beta > 2$ , define

$$\bar{G}^{(\beta)}(x) := \frac{\bar{F}(x)}{x^{\beta-\alpha}}, \text{ and } \bar{G}_I^{(\beta)}(x) := \int_x^\infty \bar{G}^{(\beta)}(u) du. \quad (25)$$

**Theorem 5.** *If the tail  $\bar{F}(\cdot)$  is regularly varying with index  $\alpha \in (1.5, 2]$ , then for any  $\beta \in (2, 2\alpha - 1)$ ,*

$$p_k = \frac{\bar{G}_I^{(\beta)}(b + n_{k-1}\mu) - \bar{G}_I^{(\beta)}(b + n_k\mu)}{\bar{G}^{(\beta)}(b)}, k \geq 1 \quad (26)$$

*yields a family of unbiased estimators ( $Z(b) = Z_K(b)/p_K : b > 0$ ) achieving*

1. *strong efficiency:  $\overline{\lim}_{b \rightarrow \infty} \frac{\text{Var}^Q[Z(b)]}{\mathbb{P}\{\tau_b < \infty\}^2} < \infty$ , and*
2. *finite expected termination time:  $\mathbb{E}^Q[\nu_b] \leq \frac{r+o(1)}{\mu(\beta-2)}b$ , as  $b \nearrow \infty$ .*

*Remark 6.* Because of the strong efficiency, we need just  $O(\epsilon^{-2}\delta^{-1})$  i.i.d. replications of  $Z(b)$  to achieve the desired relative precision. As in Remark 4, due to the bound on  $\mathbb{E}[\nu_b]$  in Theorem 5, the average computational effort for the entire estimation procedure is just  $O(\epsilon^{-2}\delta^{-1}b)$ . It is important to see this achievement in the context of Remark 5: the induced measure  $Q(\cdot)$  deviates from the zero-variance measure such that we get finite expected termination time, but only at the cost of losing vanishing relative error property to strong efficiency. Thus for the selection of  $p_k$ s as in (26), the suggested procedure ends up offering a vastly superior performance (in terms of computational complexity) compared to the zero-variance change of measure.

Given this result, it is difficult not to wonder why the tail index  $\alpha$  should be larger than 1.5 in the statement of Theorem 5, and what happens when  $\alpha \leq 1.5$ . The following result shows that it is indeed not possible to have both strong efficiency and finite expected termination time when the tail index  $\alpha < 1.5$ .

**Theorem 6.** *If the tail index  $\alpha < 1.5$ , there does not exist an assignment of  $(p_k, n_k : k \geq 1)$  such that both  $\mathbb{E}^Q[Z^2(b)]$  and  $\mathbb{E}^Q[\nu_b]$  are simultaneously finite.*

*Remark 7.* If the tail index  $\alpha = 1.5$ , the possibility of having both  $\mathbb{E}^Q[Z^2(b)]$  and  $\mathbb{E}^Q[\nu_b]$  finite will depend on the slowly varying function  $L(\cdot)$ . As we see in the proof of Proposition 6,

$$\mathbb{E}^Q[Z^2(b)]\mathbb{E}^Q[\nu_b] = \Omega \left( \int_{b^2}^\infty \sqrt{u} \bar{F}(u) du \right),$$

as  $b \nearrow \infty$ . If  $L(x) = O((\log x)^{-m})$ ,  $m \geq 2$ , the above integral is finite, whereas if  $L(x) = O(\log x)$  it is infinite; and it easily verified that the case of  $L(x) = O((\log x)^{-m})$ ,  $m \geq 2$ , goes through the proof of Theorem 5, thus achieving both strong efficiency and finite expected termination time. This illustrates the subtle dependence on the associated slowly varying function  $L(\cdot)$  for the existence of such  $p_k$ s and  $n_k$ s.

As illustrated by Theorem below, for  $\alpha \in (1, 1.5]$ , we still have algorithms that demand only  $O(b)$  units of expected computer time if we look for less stringent notions of efficiency.

**Theorem 7.** *If the tail  $\bar{F}(\cdot)$  is regularly varying with index  $\alpha \in (1, 1.5]$ , then there exists an explicit selection of  $p = (p_k : k \geq 1)$  such that the family of unbiased estimators  $(Z(b) : b > 0)$  satisfies both:*

$$\begin{aligned} \lim_{b \rightarrow \infty} \frac{\mathbb{E}^Q [Z^{1+\gamma}(b)]}{\mathbb{P}\{\tau_b < \infty\}^{1+\gamma}} &< \infty \text{ for all } \gamma \in \left(0, \frac{\alpha-1}{2-\alpha}\right), \text{ and} \\ \mathbb{E}^Q[\nu_b] &\leq Cb \text{ for some constant } C. \end{aligned} \quad (27)$$

In particular, for the following selection of  $p = (p_k : k \geq 1)$ ,

$$p_k = \frac{\bar{G}_I^{(\beta)}(b + n_{k-1}\mu) - \bar{G}_I^{(\beta)}(b + n_k\mu)}{\bar{G}^{(\beta)}(b)}, k \geq 1 \quad (28)$$

if  $\beta$  is chosen in  $(2, \alpha + \gamma^{-1}(\alpha - 1))$ , both the above inequalities are satisfied.

*Remark 8.* If the estimator  $Z(b)$  satisfies (27), similar to how we arrived at (1), it can be shown that  $O(\epsilon^{-(1+\gamma^{-1})}\delta^{-\gamma^{-1}})$  i.i.d. replications of  $Z(b)$  are enough to produce estimates having relative deviation at most  $\epsilon$  with probability at least  $1 - \delta$ . Now according to Theorem 7, the expected termination time in each replication is  $O(b)$ . Thus with the  $p_k$ s chosen as in (28), we expend just  $O(\epsilon^{-(1+\gamma^{-1})}\delta^{-\gamma^{-1}}b)$  units of computer time on an average, which is still linear in  $b$ . The price we pay by not adhering to strong efficiency is the worse dependence on the parameters  $\epsilon$  and  $\delta$ .

It is further interesting to note that a vastly different state-dependent methodology developed using Lyapunov inequalities in Blanchet and Liu (2012) also hits identical barriers and provides results similar to ours: They present algorithms that are both strongly efficient and possess  $O(b)$  expected termination time for the case of tails having index  $\alpha > 1.5$ ; whereas when  $\alpha \in (1, 1.5]$ , they provide estimators satisfying (27) along with  $O(b)$  expected termination time of a simulation run.

## 5. PROOFS OF KEY THEOREMS

**5.1. Proof of Theorem 3.** Recall that the overall estimator  $Z(b) = Z_K(b)/p_K$ , where  $p_k$  is as in (24). Second moment of the estimator  $Z(b)$  is bounded as below:

$$\begin{aligned} \mathbb{E}^Q[Z^2(b)] &= \mathbb{E}^Q \left[ \left( \frac{Z_K(b)}{p_K} \right)^2 \right] \\ &= \mathbb{E}^Q \left[ \mathbb{E}^Q \left[ \frac{Z_K^2(b)}{p_K^2} \mid K \right] \right] \\ &= \mathbb{E}^Q \left[ \mathbb{E}^Q \left[ \frac{Z_K^2(b)}{\mathbb{P}\{n_{K-1} < \tau_b \leq n_K\}^2} \cdot \frac{\mathbb{P}\{n_{K-1} < \tau_b \leq n_K\}^2}{p_K^2} \mid K \right] \right] \end{aligned} \quad (29)$$

Given  $\epsilon > 0$  and large enough  $b$ , (8) and (24) give us:

$$\frac{\mathbb{P}\{n_{K-1} < \tau_b \leq n_K\}}{p_K} \leq \frac{1 + \epsilon}{\mu} \bar{F}_I(b).$$

Also from Theorem 2, we have:

$$\mathbb{E}^Q \left[ \frac{Z_K^2(b)}{\mathbb{P}\{n_{K-1} < \tau_b \leq n_K\}^2} \mid K \right] \leq 1 + \epsilon,$$

for values of  $b$  sufficiently large. Then from (29) and (8),

$$\begin{aligned}\mathbb{E}^Q [Z^2(b)] &\leq \frac{(1+\epsilon)^3}{\mu^2} \bar{F}_I^2(b) \\ &\leq (1+\epsilon)^4 \mathbb{P}\{\tau_b < \infty\}^2,\end{aligned}$$

thus proving the asymptotically vanishing relative error property.  $\square$

**5.2. Proof of Theorem 4.** Recall that  $\nu_b$  denotes the maximum of indices of the increment random variables ( $X_i$ s) considered for simulation in a particular simulation run. From the sampling procedures in Section 4.1, it is clear that  $\nu_b \leq n_K$ . Therefore,

$$\begin{aligned}\mathbb{E}^Q[\nu_b] &\leq \sum_{k \geq 1} p_k n_k \\ &= r p_1 + \sum_{k \geq 2} r^k p_k \\ &= \frac{1}{\bar{F}_I(b)} \left( r \int_b^{b+r\mu} \bar{F}(u) du + \sum_{k \geq 1} r^{k+1} \int_{b+r^k\mu}^{b+r^{k+1}\mu} \bar{F}(u) du \right).\end{aligned}\tag{30}$$

$$\begin{aligned}\text{Since } r^k \int_{b+r^k\mu}^{b+r^{k+1}\mu} \bar{F}(u) du &= \frac{b+r^k\mu-b}{\mu} \int_{b+r^k\mu}^{b+r^{k+1}\mu} \bar{F}(u) du \\ &\leq \frac{1}{\mu} \left( \int_{b+r^k\mu}^{b+r^{k+1}\mu} u \bar{F}(u) du - b \int_{b+r^k\mu}^{b+r^{k+1}\mu} \bar{F}(u) du \right),\end{aligned}$$

$$\begin{aligned}\sum_{k \geq 1} r^{k+1} \int_{b+r^k\mu}^{b+r^{k+1}\mu} \bar{F}(u) du &\leq \frac{r}{\mu} \left( \sum_{k \geq 1} \int_{b+r^k\mu}^{b+r^{k+1}\mu} u \bar{F}(u) du - b \int_{b+r^k\mu}^{b+r^{k+1}\mu} \bar{F}(u) du \right) \\ &= \frac{r}{\mu} \left( \int_{b+r\mu}^{\infty} u \bar{F}(u) du - \int_{b+r\mu}^{\infty} \bar{F}(u) du \right) \\ &\leq \frac{r+o(1)}{\mu} \left( \frac{(b+r\mu)^2}{\alpha-2} - b \frac{b+r\mu}{\alpha-1} \right) \bar{F}(b+r\mu), \text{ as } b \nearrow \infty \\ &= \frac{r+o(1)}{\mu(\alpha-1)(\alpha-2)} b^2 \bar{F}(b),\end{aligned}\tag{31}$$

where the penultimate step follows from Karamata's theorem (see (10)), and the final step just uses long-tailed nature of  $\bar{F}(\cdot)$ . Also note that:  $\int_b^{b+r\mu} \bar{F}(u) du \leq r\mu \bar{F}(b)$ , and by application of Karamata's theorem, we have  $\bar{F}_I(b) \sim \frac{b\bar{F}(b)}{\alpha-1}$ , as  $b \nearrow \infty$ . Therefore from (30),

$$\mathbb{E}^Q[\nu_b] \leq \frac{r+o(1)}{\mu(\alpha-2)} b, \text{ as } b \nearrow \infty,$$

thus yielding the required bound on the expected termination time.  $\square$

*Remark 9.* Similar to how we arrived at (31), lower bounds can be obtained to show that  $\mathbb{E}^Q[\nu_b] = \Omega\left(\int_b^{\infty} u \bar{F}(u) du\right)$ . If the tail index  $\alpha < 2$ ,  $\int_b^{\infty} u \bar{F}(u) du$  turns out to be infinite, and subsequently  $\mathbb{E}^Q[\nu_b] = \infty$ . Though the assignment of  $p_k$ s in (24) yields vanishing relative error

for any  $\alpha > 1$ , it fails to provide algorithms which have finite expected termination time when the increment random variables  $X$  have infinite variance (e.g., when  $\alpha < 2$ ), thus making this choice of  $p_k$  not suitable for practice.

### 5.3. Proof of Theorem 5.

1. *Variance of  $Z(b)$* : Since  $Q(K = k) = p_k$ ,

$$\begin{aligned}\mathbb{E}^Q[Z^2(b)] &= \mathbb{E}^Q\left[\frac{Z_K^2(b)}{p_K^2}\right] = \sum_k p_k \frac{\mathbb{E}^Q[Z_k^2(b)]}{p_k^2} \\ &= \sum_k \frac{\mathbb{E}^Q[Z_k^2(b)]}{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}^2} \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}^2}{p_k}.\end{aligned}$$

Thanks to the uniformly efficient estimators developed in Section 4.1, Theorem 2 helps us to write:

$$\mathbb{E}^Q[Z^2(b)] \leq (1 + \epsilon)^2 \sum_k \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}^2}{p_k}, \quad (32)$$

for large values of  $b$ . Due to the uniform convergence in (8), and the assignment of  $p_k$ s as in (26), we can write,

$$\frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}}{p_k} \leq (1 + \epsilon) \frac{\bar{F}_I(b + n_{k-1}\mu) - \bar{F}_I(b + n_k\mu)}{\bar{G}_I^{(\beta)}(b + n_{k-1}\mu) - \bar{G}_I^{(\beta)}(b + n_k\mu)} \bar{G}_I^{(\beta)}(b),$$

uniformly in  $k$ . Note that,

$$\begin{aligned}\bar{F}_I(b + n_{k-1}\mu) - \bar{F}_I(b + n_k\mu) &= \int_{b+n_{k-1}\mu}^{b+n_k\mu} \bar{F}(u) du \leq (n_k - n_{k-1})\mu \bar{F}(b + n_{k-1}\mu), \\ \bar{G}_I^{(\beta)}(b + n_{k-1}\mu) - \bar{G}_I^{(\beta)}(b + n_k\mu) &= \int_{b+n_{k-1}\mu}^{b+n_k\mu} \bar{G}^{(\beta)}(u) du \geq (n_k - n_{k-1})\mu \bar{G}^{(\beta)}(b + n_k\mu), \text{ and} \\ \frac{\bar{G}^{(\beta)}(b + n_{k-1}\mu)}{\bar{G}^{(\beta)}(b + n_k\mu)} &\leq \frac{\bar{G}^{(\beta)}(b + n_{k-1}\mu)}{\bar{G}^{(\beta)}(r(b + n_{k-1}\mu))} \leq (1 + \epsilon)r^\beta,\end{aligned}$$

due to the regularly varying nature of  $\bar{G}^{(\beta)}(\cdot)$ . Therefore for values of  $b$  sufficiently large,

$$\begin{aligned}\frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}}{p_k} &\leq (1 + \epsilon) \frac{\bar{F}(b + n_{k-1}\mu)}{\bar{G}^{(\beta)}(b + n_k\mu)} \frac{\bar{G}^{(\beta)}(b + n_k\mu)}{\bar{G}^{(\beta)}(b + n_{k-1}\mu)} \bar{G}_I^{(\beta)}(b) \\ &= (1 + \epsilon)^2 r^\beta (b + n_{k-1}\mu)^{\beta-\alpha} \bar{G}_I^{(\beta)}(b),\end{aligned} \quad (33)$$

for all  $k$ , because  $\frac{\bar{F}(x)}{\bar{G}^{(\beta)}(x)} = x^{\beta-\alpha}$ . Then from (32)

$$\begin{aligned}\mathbb{E}^Q[Z^2(b)] &\leq (1 + \epsilon)^4 r^\beta \bar{G}_I^{(\beta)}(b) \sum_k (b + n_{k-1}\mu)^{\beta-\alpha} \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\} \\ &\leq (1 + \epsilon)^5 r^\beta \bar{G}_I^{(\beta)}(b) \sum_k (b + n_{k-1}\mu)^{\beta-\alpha} \int_{b+n_{k-1}\mu}^{b+n_k\mu} \bar{F}(u) du,\end{aligned}$$



because of (8). Consequently,

$$\begin{aligned}\mathbb{E}^Q[Z^2(b)] &\leq (1+\epsilon)^5 r^\beta \bar{G}_I^{(\beta)}(b) \sum_k \int_{b+n_{k-1}\mu}^{b+n_k\mu} u^{\beta-\alpha} \bar{F}(u) du \\ &\leq (1+\epsilon)^5 r^\beta \bar{G}_I^{(\beta)}(b) \int_b^\infty u^{\beta-\alpha} \bar{F}(u) du \\ &\leq (1+\epsilon)^6 r^\beta \bar{G}_I^{(\beta)}(b) b^{\beta-\alpha+1} \frac{\bar{F}(b)}{2\alpha-\beta-1},\end{aligned}$$

for large enough values of  $b$  due to Karamata's theorem if  $2\alpha - \beta > 1$ . This is indeed true because  $\beta$  is assumed smaller than  $2\alpha - 1$  in the statement of Theorem 5. Further  $(\alpha - 1)\bar{F}_I(b) \sim b\bar{F}(b)$  and  $b^{\beta-\alpha}\bar{G}_I^{(\beta)}(b) \sim \bar{F}_I(b)$ , as  $b \nearrow \infty$ . Therefore,

$$\lim_{b \rightarrow \infty} \frac{\mathbb{E}^Q[Z^2(b)]}{\bar{F}_I^2(b)} \leq \frac{(\alpha - 1)r^\beta}{2\alpha - \beta - 1} < \infty.$$

Now since  $\mathbb{P}\{\tau_b < \infty\} \sim \mu^{-1}\bar{F}_I(b)$ , we have strong efficiency.

2. *Expected termination time:* Since  $\nu_b \leq n_K$ ,  $\mathbb{E}^Q[\nu_b] \leq \mathbb{E}^Q[n_K] = \sum_k p_k n_k$ . For the choice of  $p_k$  in (26), following exactly the same steps in the proof of Theorem 4, we arrive at:

$$\mathbb{E}^Q[\nu_b] \leq \frac{r}{\mu} \left( \mu \int_b^{b+r\mu} \bar{G}^{(\beta)}(u) du + \int_{b+r\mu}^\infty u \bar{G}^{(\beta)}(u) du - b \int_{b+r\mu}^\infty \bar{G}^{(\beta)}(u) du \right).$$

Since  $\bar{G}^{(\beta)}(\cdot)$  is regularly varying with tail index larger than 2, by application of Karamata's theorem, we have:

$$\int_{b+r\mu}^\infty u \bar{G}^{(\beta)}(u) du \sim \frac{(b+r\mu)^2}{\beta-2} \bar{G}^{(\beta)}(b+r\mu),$$

which would not have been the case if we had persisted with using  $\bar{F}_I(\cdot)$  instead of  $\bar{G}_I^{(\beta)}(\cdot)$  for  $p_k$ . Again following the remaining steps in the proof of Theorem 4, we conclude that:

$$\mathbb{E}^Q[\nu_b] \leq \frac{r + o(1)}{\mu(\beta-2)} b, \text{ as } b \nearrow \infty,$$

thus yielding finite termination time even when the zero-variance measure fails to offer this desirable property.  $\square$

**5.4. Proof of Theorem 6.** Since  $Q(K = k) = p_k$ , see that:

$$\mathbb{E}^Q[Z^2(b)] = \mathbb{E}^Q \left[ \frac{Z_K^2(b)}{p_K^2} \right] = \sum_k \frac{\mathbb{E}^Q[Z_k^2(b)]}{p_k} \geq \sum_k \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}^2}{p_k},$$

because of Jensen's inequality. To arrive at a contradiction, let us assume that both  $\mathbb{E}^Q[Z^2(b)]$  and  $\mathbb{E}^Q[\nu_b]$  are finite. Then,

$$\begin{aligned}\mathbb{E}^Q[Z^2(b)] \mathbb{E}^Q[\nu_b] &\geq \left( \sum_k \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}^2}{p_k} \right) \left( \sum_k p_k n_k \right) \\ &\geq \left( \sum_k \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}}{\sqrt{p_k}} \cdot \sqrt{p_k n_k} \right)^2 \\ &= \left( \sum_k \sqrt{n_k} \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\} \right)^2.\end{aligned}\tag{34}$$

where the penultimate step follows from Cauchy-Schwarz inequality. Due to the uniform convergence in (8), given  $\epsilon > 0$ , for  $b$  large enough:

$$\begin{aligned} \sum_k \sqrt{n_k} \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\} &\geq (1 - \epsilon) \sum_k \sqrt{n_k} \int_{n_{k-1}\mu}^{n_k\mu} \bar{F}(b+u) du \\ &\geq \frac{1 - \epsilon}{\sqrt{\mu}} \sum_k \int_{n_{k-1}\mu}^{n_k\mu} \sqrt{u} \bar{F}(b+u) du \\ &= \frac{1 - \epsilon}{\sqrt{\mu}} \int_0^\infty \sqrt{u} \bar{F}(b+u) du. \end{aligned}$$

Now it can be seen easily that the RHS is finite only when  $\alpha \geq 1.5$ , via the following change of variable and the subsequent integration of the resulting regularly varying tail:

$$\begin{aligned} \int_0^\infty \sqrt{u} \bar{F}(b+u) du &= \int_b^\infty \sqrt{u-b} \bar{F}(u) du \\ &\geq \int_{b^2}^\infty \sqrt{u} \cdot \sqrt{1 - \frac{b}{u}} \bar{F}(u) du \\ &\geq \sqrt{1 - \frac{1}{b}} \int_{b^2}^\infty \sqrt{u} \bar{F}(u) du, \end{aligned}$$

which cannot be finite if  $\alpha < 1.5$ , thus arriving at the desired contradiction. Therefore from (34), we conclude that we cannot have both the second moment of  $Z(b)$  and the expected termination time  $\mathbb{E}^Q[\nu_b]$  to be simultaneously finite if the tail index  $\alpha < 1.5$ .  $\square$

**5.5. Proof of Theorem 7.** The proof is similar to that of Theorem 5, and we provide only an outline of the steps involved. Since  $Q(K = k) = p_k$ , as in (32),

$$\mathbb{E}^Q[Z^{1+\gamma}(b)] \leq (1 + \epsilon)^{1+\gamma} \sum_k \left( \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}}{p_k} \right)^\gamma \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\},$$

for sufficiently large values of  $b$ . Then using (33) and (8),

$$\begin{aligned} \mathbb{E}^Q[Z^{1+\gamma}(b)] &\leq (1 + \epsilon)^{1+3\gamma} \left( \bar{G}_I^{(\beta)}(b) \right)^\gamma \sum_k (b + n_{k-1}\mu)^{\gamma(\beta-\alpha)} \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\} \\ &\leq (1 + \epsilon)^{2+3\gamma} \left( \bar{G}_I^{(\beta)}(b) \right)^\gamma \int_b^\infty u^{\gamma(\beta-\alpha)} \bar{F}(u) du, \end{aligned}$$

which follows from the routine calculation in the proof of Theorem 5. Since  $\beta$  is smaller than  $\alpha + \gamma^{-1}(\alpha - 1)$  as in the statement of Theorem 7, the tail index of the integrand,  $\alpha - \gamma(\beta - \alpha) > 1$ . Therefore we can apply Karamata's theorem to conclude that for values of  $b$  large enough,

$$\mathbb{E}^Q[Z^{1+\gamma}(b)] \leq (1 + \epsilon)^{3+3\gamma} \left( \bar{G}_I^{(\beta)}(b) \right)^\gamma \frac{b^{\gamma(\beta-\alpha)+1}}{\alpha - \gamma(\beta - \alpha) - 1} \bar{F}(b).$$

Now observing that  $(\alpha - 1)\bar{F}_I(b) \sim b\bar{F}(b)$ ,  $b^{\beta-\alpha}\bar{G}_I^{(\beta)}(b) \sim \bar{F}_I(b)$ , and  $\mathbb{P}\{\tau_b < \infty\} \sim \mu^{-1}\bar{F}_I(b)$  as  $b \nearrow \infty$ , we have:

$$\overline{\lim}_{b \rightarrow \infty} \frac{\mathbb{E}^Q[Z^{1+\gamma}(b)]}{\mathbb{P}\{\tau_b < \infty\}^{1+\gamma}} \leq \frac{\mu^2(\alpha - 1)}{\alpha - \gamma(\beta - \alpha) - 1} < \infty.$$

Since  $\beta$  is ensured to be larger than 2, the same proof for  $\mathbb{E}^Q[\nu_b] = O(b)$  goes through.  $\square$

## 6. NUMERICAL EXPERIMENTS

In this section, we present the results of numerical simulation experiments performed on examples previously considered in literature, and compare the performance of our algorithms.

**6.1. Example 1 - estimation of  $\mathbb{P}\{S_n > b\}$ .** Take  $X = \Lambda R$ , where  $\mathbb{P}\{\Lambda > x\} = 1 \wedge x^{-4}$ ,  $R \sim \text{Laplace}(1)$ , and  $\Lambda$  is independent of  $R$ . We use  $N = 10,000$  simulation runs to estimate  $\mathbb{P}\{S_n > n\}$  for  $n = 100, 500$  and  $1000$ . In Table 1, we compare the numerical estimates obtained by our simulation procedure with the true values of  $\mathbb{P}\{S_n > n\}$  evaluated in Blanchet and Liu (2008) via inverse transform techniques; further, a comparison of performance of our methodology with Algorithms 1 and 2 in Blanchet and Liu (2008) (referred to as BL1 and BL2) has also been presented. From the columns CV, CV of BL1, and CV of BL2, it can be inferred that our state-independent simulation procedures yield estimators with substantially lower coefficient of variation throughout the range of values considered.

n	$\mathbb{P}\{S_n > n\}$	Estimate ( $\hat{z}$ ) for $\mathbb{P}\{S_n > n\}$	Std. error	CV of $\hat{z}$	CV of BL1	CV of BL2
100	$2.21 \times 10^{-5}$	$2.17 \times 10^{-5}$	$4.31 \times 10^{-7}$	1.97	10.3	4.7
500	$1.04 \times 10^{-7}$	$1.05 \times 10^{-7}$	$6.91 \times 10^{-10}$	0.66	1.0	4.1
1000	$1.25 \times 10^{-8}$	$1.29 \times 10^{-8}$	$6.91 \times 10^{-11}$	0.53	1.1	3.8

Table 1: Numerical result for Example 1 - here Std. error denotes the standard deviation of the estimator of  $\mathbb{P}\{S_n > n\}$  based on 10,000 simulation runs; CV denotes the empirically observed coefficient of variation

**6.2. Example 2 - estimation of  $\mathbb{P}\{\tau_b < \infty\}$ .** To facilitate comparison with existing methods, we use the following example from Blanchet and Glynn (2008): Consider an M/G/1 queue with traffic intensity  $\rho = 0.5$  and Pareto service times having tail  $\mathbb{P}\{V > t\} = (1 + t)^{-2.5}$ . The aim is to estimate the probability that this queue develops a waiting time  $b$  in stationarity by equivalently estimating the level crossing probabilities  $\mathbb{P}\{\tau_b < \infty\}$  of the associated negative drift random walk. For this example, we use the simulation procedures discussed in Section 4 and compare the results with that of the existing algorithms in literature in Table 2. While Algorithms AK (in Asmussen and Kroese (2006)) and DLW (in Dupuis et al. (2007)) restrict the arrivals to be Poisson, the schemes BGL, BG and BL referring to the algorithms, respectively, in Blanchet et al. (2007), Blanchet and Glynn (2008) and Blanchet and Liu (2012) do not impose any such restriction.

In our implementation,  $r$  has been chosen to be 2 to keep the expected termination time low, as suggested by Theorem 4. The results reported in Table 2 correspond to the simulation estimates of  $\mathbb{P}\{\tau_b < \infty\}$  for values of  $b = 10^2, 10^3$  and  $10^4$  using  $N = 10,000$  simulation runs. From Table 2, it can be inferred that the error offered by the estimates of our simpler state-independent procedure is much smaller when compared with other existing algorithms. Table 3 gives a comparison of coefficient of variation of the estimators empirically observed for different values of  $r$ , and a fixed  $b = 10^3$ . It can be seen from Table 3 as well that choosing  $r = 2$  helps in keeping the relative error low.

## 7. CONCLUSION

In this paper we revisited the problem of efficient simulation of commonly encountered rare event probabilities associated with random walks having regularly varying heavy-tailed increments.

Estimation Std. error CV	$b = 10^2$	$b = 10^3$	$b = 10^4$
	$9.75 \times 10^{-4}$	$3.15 \times 10^{-5}$	$9.98 \times 10^{-7}$
Proposed	$4.11 \times 10^{-6}$	$7.89 \times 10^{-8}$	$1.39 \times 10^{-9}$
method	0.42	0.25	0.14
	$1.20 \times 10^{-3}$	$3.15 \times 10^{-5}$	$9.98 \times 10^{-7}$
AK	$1.48 \times 10^{-5}$	$2.19 \times 10^{-7}$	$6.95 \times 10^{-9}$
	1.23	0.70	0.70
	$1.05 \times 10^{-3}$	$3.16 \times 10^{-5}$	$9.91 \times 10^{-7}$
DLW	$5.20 \times 10^{-6}$	$1.69 \times 10^{-7}$	$2.99 \times 10^{-9}$
	0.50	0.53	0.30
	$1.02 \times 10^{-3}$	$3.17 \times 10^{-5}$	$1.13 \times 10^{-6}$
BGL	$3.84 \times 10^{-5}$	$1.60 \times 10^{-6}$	$7.28 \times 10^{-8}$
	3.76	5.05	6.44
	$1.08 \times 10^{-3}$	$3.15 \times 10^{-5}$	$9.98 \times 10^{-7}$
BG	$5.97 \times 10^{-6}$	$9.73 \times 10^{-8}$	$2.07 \times 10^{-9}$
	0.55	0.31	0.21
	$1.05 \times 10^{-3}$	$3.18 \times 10^{-5}$	$9.88 \times 10^{-7}$
BL	$3.76 \times 10^{-5}$	$2.60 \times 10^{-7}$	$8.19 \times 10^{-9}$
	3.58	0.82	0.83

Table 2: Numerical result for Example 2 - here Std. error denotes the standard deviation of the estimator of  $\mathbb{P}\{\tau_b < \infty\}$  based on 10,000 simulation runs; CV denotes the empirically observed coefficient of variation

$r$	Estimate	Std. error	CV
2	$3.15 \times 10^{-5}$	$7.89 \times 10^{-8}$	0.25
10	$3.16 \times 10^{-5}$	$1.03 \times 10^{-7}$	0.33
100	$3.16 \times 10^{-5}$	$1.55 \times 10^{-7}$	0.49

Table 3: Comparison of relative errors for different choices of  $r$  in Example 2 with  $b = 1000$ ; here Std. error denotes the standard deviation of the estimator of  $\mathbb{P}\{\tau_b < \infty\}$  based on 10,000 simulation runs; CV denotes the empirically observed coefficient of variation

These comprised the large deviations probability of a random walk exceeding large values as well as the level crossing probability of a negative-drift random walk. In the existing literature there are results that suggest that state-independent methods for such probabilities are difficult to design. Significant research over the last few years has resulted in sophisticated state-dependent importance sampling techniques for estimating these probabilities. Our key contribution has been to challenge this view by showing that simple state-independent importance sampling methods, that are at least as efficient as the existing state-dependent methods, can indeed be devised to estimate these probabilities.

Our approach relied on partitioning the rare event of interest into elementary events that were amenable to straight forward state-independent importance sampling methods. We expect that this approach will generalize to more complex, multi-dimensional problems, and for similar problems involving Weibull-type sub-exponential tail distributions.

## REFERENCES

- R. J. Adler, R. E. Feldman, and M. S. Taqqu, editors. *A practical guide to heavy tails*. Birkhäuser Boston Inc., Boston, MA, 1998. ISBN 0-8176-3951-9. Statistical techniques and applications, Papers from the workshop held in Santa Barbara, CA, December 1995.
- A. Agarwal, S. Juneja, and S. Dey. Efficient simulation of large deviations events for sums of random vectors using saddle point representations. *Journal of Applied Probability*, 2013.
- S. Asmussen and P. W. Glynn. *Stochastic simulation: algorithms and analysis*, volume 57 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007. ISBN 978-0-387-30679-7.
- S. Asmussen and C. Kluppelberg. Large deviations results for subexponential tails, with applications to insurance risk. *Stochastic Processes and their Applications*, 64(1):103 – 125, 1996. ISSN 0304-4149. doi: 10.1016/S0304-4149(96)00087-7. URL <http://www.sciencedirect.com/science/article/pii/S0304414996000877>.
- S. Asmussen and D. P. Kroese. Improved algorithms for rare event simulation with heavy tails. *Adv. in Appl. Probab.*, 38(2):545–558, 2006. ISSN 0001-8678. doi: 10.1239/aap/1151337084. URL <http://dx.doi.org/10.1239/aap/1151337084>.
- S. Asmussen, K. Binswanger, and B. Højgaard. Rare events simulation for heavy-tailed distributions. *Bernoulli*, 6(2):303–322, 2000. ISSN 1350-7265. doi: 10.2307/3318578. URL <http://dx.doi.org/10.2307/3318578>.
- A. Bassamboo, S. Juneja, and A. Zeevi. On the inefficiency of state-independent importance sampling in the presence of heavy tails. *Oper. Res. Lett.*, 35(2):251–260, 2007. ISSN 0167-6377. doi: 10.1016/j.orl.2006.02.002. URL <http://dx.doi.org/10.1016/j.orl.2006.02.002>.
- A. Bassamboo, S. Juneja, and A. J. Zeevi. Portfolio credit risk with extremal dependence: Asymptotic analysis and efficient simulation. *Operations Research*, 56(3):593–606, 2008.
- J. Blanchet and P. Glynn. Efficient rare-event simulation for the maximum of heavy-tailed random walks. *Ann. Appl. Probab.*, 18(4):1351–1378, 2008. ISSN 1050-5164. doi: 10.1214/07-AAP485. URL <http://dx.doi.org/10.1214/07-AAP485>.
- J. Blanchet and J. Liu. Efficient simulation and conditional functional limit theorems for ruinous heavy-tailed random walks. *Stochastic Processes and their Applications*, 122(8):2994 – 3031, 2012. ISSN 0304-4149. doi: 10.1016/j.spa.2012.05.001. URL <http://www.sciencedirect.com/science/article/pii/S0304414912000877>.
- J. Blanchet, P. Glynn, and J. Liu. Fluid heuristics, lyapunov bounds and efficient importance sampling for a heavy-tailed g/g/1 queue. *Queueing Systems*, 57(2-3):99–113, 2007. ISSN 0257-0130. doi: 10.1007/s11134-007-9047-4. URL <http://dx.doi.org/10.1007/s11134-007-9047-4>.
- J. H. Blanchet and J. Liu. State-dependent importance sampling for regularly varying random walks. *Adv. in Appl. Probab.*, 40(4):1104–1128, 2008. ISSN 0001-8678. URL <http://projecteuclid.org/getRecord?id=euclid.aap/1231340166>.
- J. H. Blanchet, K. Leder, and P. W. Glynn. Efficient simulation of light-tailed sums: an old-folk song sung to a faster new tune... In *Monte Carlo and quasi-Monte Carlo methods 2008*, pages 227–248. Springer, Berlin, 2009. doi: 10.1007/978-3-642-04107-5\_13. URL [http://dx.doi.org/10.1007/978-3-642-04107-5\\_13](http://dx.doi.org/10.1007/978-3-642-04107-5_13).

- A. A. Borovkov and K. A. Borovkov. *Asymptotic analysis of random walks*, volume 118 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2008. ISBN 978-0-521-88117-3. doi: 10.1017/CBO9780511721397. URL <http://dx.doi.org/10.1017/CBO9780511721397>. Heavy-tailed distributions, Translated from the Russian by O. B. Borovkova.
- H. Chan, S. Deng, and T. Lai. Rare-event simulation of heavy-tailed random walks by sequential importance sampling and resampling. *Preprint*, 2012.
- A. B. Dieker and M. Mandjes. On asymptotically efficient simulation of large deviation probabilities. *Adv. in Appl. Probab.*, 37(2):539–552, 2005. ISSN 0001-8678. doi: 10.1239/aap/1118858638. URL <http://dx.doi.org/10.1239/aap/1118858638>.
- P. Dupuis and H. Wang. Importance sampling, large deviations, and differential games. *Stoch. Stoch. Rep.*, 76(6):481–508, 2004. ISSN 1045-1129. doi: 10.1080/10451120410001733845. URL <http://dx.doi.org/10.1080/10451120410001733845>.
- P. Dupuis, K. Leder, and H. Wang. Importance sampling for sums of random variables with regularly varying tails. *ACM Trans. Model. Comput. Simul.*, 17(3), July 2007. ISSN 1049-3301. doi: 10.1145/1243991.1243995. URL <http://doi.acm.org/10.1145/1243991.1243995>.
- P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*, volume 33 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. ISBN 3-540-60931-8. For insurance and finance.
- S. Foss, D. Korshunov, and S. Zachary. *An introduction to heavy-tailed and subexponential distributions*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2011. ISBN 978-1-4419-9472-1. doi: 10.1007/978-1-4419-9473-8. URL <http://dx.doi.org/10.1007/978-1-4419-9473-8>.
- P. Glasserman and J. Li. Importance sampling for portfolio credit risk. *Management Science*, 51(11):1643–1656, 2005.
- P. W. Glynn and W. Whitt. The asymptotic efficiency of simulation estimators. *Oper. Res.*, 40(3):505–520, 1992. ISSN 0030-364X. doi: 10.1287/opre.40.3.505. URL <http://dx.doi.org/10.1287/opre.40.3.505>.
- J. M. Hammersley and D. C. Handscomb. *Monte Carlo methods*. Methuen & Co. Ltd., London, 1965.
- S. Juneja and P. Shahabuddin. Simulating heavy tailed processes using delayed hazard rate twisting. *ACM Trans. Model. Comput. Simul.*, 12(2):94–118, Apr. 2002. ISSN 1049-3301. doi: 10.1145/566392.566394. URL <http://doi.acm.org/10.1145/566392.566394>.
- T. Mikosch and A. V. Nagaev. Large deviations of heavy-tailed sums with applications in insurance. *Extremes*, 1(1):81–110, 1998. ISSN 1386-1999. doi: 10.1023/A:1009913901219. URL <http://dx.doi.org/10.1023/A:1009913901219>.
- S. Parekh and J. Walrand. A quick simulation method for excessive backlogs in networks of queues. *IEEE Trans. Automat. Control*, 34(1):54–66, 1989. ISSN 0018-9286. doi: 10.1109/9.8649. URL <http://dx.doi.org/10.1109/9.8649>.



- A. M. K. Rajhaa and S. Juneja. State-independent importance sampling for estimating large deviation probabilities in heavy-tailed random walks. In *Performance Evaluation Methodologies and Tools (VALUETOOLS), 2012 6th International Conference on*, pages 127–135, oct. 2012.
- S. I. Resnick. Heavy tail modeling and teletraffic data. *Ann. Statist.*, 25(5):1805–1869, 1997. ISSN 0090-5364. doi: 10.1214/aos/1069362376. URL <http://dx.doi.org/10.1214/aos/1069362376>. With discussion and a rejoinder by the author.
- J. S. Sadowsky. On Monte Carlo estimation of large deviations probabilities. *Ann. Appl. Probab.*, 6(2):399–422, 1996. ISSN 1050-5164. doi: 10.1214/aoap/1034968137. URL <http://dx.doi.org/10.1214/aoap/1034968137>.
- J. S. Sadowsky and J. A. Bucklew. On large deviations theory and asymptotically efficient Monte Carlo estimation. *IEEE Trans. Inform. Theory*, 36(3):579–588, 1990. ISSN 0018-9448. doi: 10.1109/18.54903. URL <http://dx.doi.org/10.1109/18.54903>.
- D. Siegmund. Importance sampling in the Monte Carlo study of sequential tests. *Ann. Statist.*, 4(4):673–684, 1976. ISSN 0090-5364.

## APPENDIX

Here we present proofs of Lemmas 1, 2, 4 and 5. To prove Lemmas 1 and 4, we need Lemmas 8 and 9, which are stated and proved below. The proof of Lemma 8 follow the lines of Borovkov and Borovkov (2008), where bounds for similar integrals have been derived.

**Lemma 8.** *For any pair of sequences  $\{x_n\}, \{\phi_n\}$  satisfying  $x_n \nearrow \infty$  and  $\phi_n x_n \nearrow \infty$ , the integral,*

$$\int_{-\infty}^{x_n} e^{\phi_n x} F(dx) \leq 1 + c\phi_n^\kappa + e^{2\alpha} \bar{F}\left(\frac{2\alpha}{\phi_n}\right) + e^{\phi_n x_n} \bar{F}(x_n)(1 + o(1)), \text{ as } n \nearrow \infty,$$

for any  $0 < \kappa < \alpha \wedge 2$ , and some constant  $c$  which does not depend on  $n$  and  $b$ .

*Proof.* We split the region of integration into  $(-\infty, \gamma/\phi_n]$  and  $(\gamma/\phi_n, x_n]$  for some constant  $\gamma > 0$ ; the partition is such that the integrand stays bounded in the former despite its growth to  $(-\infty, \infty)$ . Let  $I_1 := \int_{-\infty}^{\gamma/\phi_n} e^{\phi_n x} F(dx)$  and  $I_2 := \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} F(dx)$ . Since  $e^{\phi_n x} \leq 1 + \phi_n x + \phi_n^\kappa |x|^\kappa e^{\phi_n x}$ ,

$$\begin{aligned} I_1 &\leq \int_{-\infty}^{\gamma/\phi_n} F(dx) + \phi_n \int_{-\infty}^{\gamma/\phi_n} x F(dx) + \phi_n^\kappa \int_{-\infty}^{\gamma/\phi_n} |x|^\kappa e^{\phi_n x} F(dx) \\ &\leq \int_{-\infty}^{\infty} F(dx) + \phi_n \int_{-\infty}^{\infty} x F(dx) + \phi_n^\kappa e^\gamma \int_{-\infty}^{\infty} |x|^\kappa F(dx) \\ &= 1 + c\phi_n^\kappa, \end{aligned} \tag{35}$$

where  $c := e^\gamma \int_{-\infty}^{\infty} |x|^\kappa F(dx) < \infty$  because  $\mathbb{E}|X|^\kappa < \infty$ ; this follows because  $\kappa < \alpha$  and from Assumption 1. We have also used  $\mathbb{E}X = 0$  to arrive at (35). Integrating by parts for the second

integral  $I_2$  :

$$\begin{aligned}
I_2 &= - \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} \bar{F}(dx) \\
&= e^{\phi_n \gamma/\phi_n} \bar{F}(\gamma/\phi_n) - e^{\phi_n x_n} \bar{F}(x_n) + \phi_n \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} \bar{F}(x) dx \\
&\leq e^\gamma \bar{F}(\gamma/\phi_n) + I'_2,
\end{aligned} \tag{36}$$

where,  $I'_2 := \phi_n \int_{\gamma/\phi_n}^{x_n} e^{\phi_n x} \bar{F}(x) dx$ . Now the change of variable  $u = \phi_n(x_n - x)$  results in:

$$\begin{aligned}
I'_2 &= e^{\phi_n x_n} \int_0^{\phi_n x_n - \gamma} e^{-u} \bar{F}\left(x_n - \frac{u}{\phi_n}\right) du \\
&= e^{\phi_n x_n} \bar{F}(x_n) \int_0^{\phi_n x_n - \gamma} e^{-u} g_n(u) du,
\end{aligned} \tag{37}$$

where,

$$g_n(u) := \frac{\bar{F}\left(x_n - \frac{u}{\phi_n}\right)}{\bar{F}(x_n)} = \frac{\bar{F}\left(x_n \left(1 - \frac{u}{\phi_n x_n}\right)\right)}{\bar{F}(x_n)}.$$

Since  $L(\cdot)$  is slowly varying and  $\phi_n x_n \rightarrow \infty$ , given any  $\delta > 0$ , for all  $n$  large enough we have:

$$(1 - \delta) \left(1 - \frac{u}{\phi_n x_n}\right)^{-\alpha + \delta} \leq g_n(u) \leq (1 + \delta) \left(1 - \frac{u}{\phi_n x_n}\right)^{-\alpha - \delta}.$$

This preliminary fact about slowly varying functions can be found in, e.g., Theorem 1.1.4 of Borovkov and Borovkov (2008). So for any fixed  $u$ , we have  $g_n(u) \rightarrow 1$  as  $n \nearrow \infty$ . Now fix  $\delta = \frac{\alpha}{2}$ . Then for  $n$  large enough,

$$g_n(u) \leq \left(1 + \frac{\alpha}{2}\right) \left(1 - \frac{u}{\phi_n x_n}\right)^{-\frac{3\alpha}{2}}. \tag{38}$$

Let  $h(u) = (1 - u/\phi_n x_n)^{-\frac{3\alpha}{2}}$ . Since  $\log h(0) = 0$  and  $\frac{d}{du}(\log(h(u))) \leq \frac{3\alpha}{2\gamma}$  for  $0 \leq u \leq \phi_n x_n - \gamma$ , we have  $h(u) \leq e^{\frac{3\alpha u}{2\gamma}}$  on the same interval. Therefore if we choose  $\gamma = 2\alpha$ , the integrand in  $I'_2$  is bounded for large enough  $n$  by an integrable function as below:

$$\left| e^{-u} g_n(u) \mathbf{1}_{\{0 \leq u \leq \phi_n x_n - \gamma\}} \right| \leq \left| e^{-u} \left(1 + \frac{\alpha}{2}\right) h(u) \mathbf{1}_{\{0 \leq u \leq \phi_n x_n - \gamma\}} \right| \leq \left(1 + \frac{\alpha}{2}\right) e^{-u + \frac{3\alpha u}{2\gamma}} = \left(1 + \frac{\alpha}{2}\right) e^{-\frac{u}{4}}.$$

Applying dominated convergence theorem, we get

$$\int_0^{\phi_n x_n - \gamma} e^{-u} g_n(u) du \sim 1 \text{ as } n \nearrow \infty.$$

Since  $\int_{-\infty}^{x_n} e^{\phi_n x} F(dx) = I_1 + I_2$ , combining this result with (35), (36) and (37), completes the proof.  $\square$

**Lemma 9.** *Given any  $\epsilon > 0$ , uniformly for  $b > n^{\beta + \epsilon}$ , we have:*

(a)  $n\theta_{n,b}^\kappa \searrow 0$  for some  $0 < \kappa < \alpha$ , and

(b)  $\bar{F}\left(\frac{2\alpha}{\theta_{n,b}}\right) = O\left(\frac{1}{n}\right)$ , as  $n \nearrow \infty$ .

*Proof.* (a) We have  $\bar{F}(x) = \frac{L(x)}{x^\alpha}$ . Since  $L(\cdot)$  is slowly varying, given any  $\delta > 0$  for sufficiently large values of  $b$ , we have  $b^{-\delta} \leq L(b) \leq b^\delta$ , thus yielding  $L(b) = b^{o(1)}$  as  $b \nearrow \infty$ . Further noting that  $b > n^{\beta+\epsilon}$  helps us to write:

$$n\theta_{n,b}^\kappa = \frac{n}{b^\kappa} \log^\kappa \left( \frac{1}{n\bar{F}(b)} \right) \leq n^{1-\kappa(\beta+\epsilon)} \log^\kappa \left( \frac{b^\alpha}{nL(b)} \right).$$

If we take,

$$\kappa := \begin{cases} 2, & \text{if } \alpha > 2 \\ (1+\epsilon)/(\frac{1}{\alpha}+\epsilon), & \text{if } 1 < \alpha \leq 2 \end{cases} \quad (39)$$

then  $\kappa < \alpha$ , and  $\kappa(\beta+\epsilon) \geq 1 + \epsilon/2$ . Then  $n\theta_{n,b}^\kappa \searrow 0$  as  $n \nearrow \infty$ , uniformly for  $b > n^{\beta+\epsilon}$ .

(b) We have  $\theta_{n,b} := -\log(n\bar{F}(b))/b$ . Therefore,

$$n\bar{F}\left(\frac{2\alpha}{\theta_n}\right) = n\bar{F}(b) \frac{\bar{F}\left(\frac{2\alpha b}{-\log(n\bar{F}(b))}\right)}{\bar{F}(b)}.$$

Since  $\bar{F}(\cdot)$  is regularly varying, given any  $\delta > 0$ , for  $n$  large enough,

$$\frac{\bar{F}\left(\frac{2\alpha b}{-\log(n\bar{F}(b))}\right)}{\bar{F}(b)} \leq \left( \frac{-\log(n\bar{F}(b))}{2\alpha} \right)^{\alpha+\delta}.$$

The above inequality is just an application of Theorem 1.1.4 of Borovkov and Borovkov (2008). Therefore,

$$n\bar{F}\left(\frac{2\alpha}{\theta_n}\right) \leq n \frac{L(b)}{b^\alpha} \left( \frac{-\log(n\bar{F}(b))}{2\alpha} \right)^{\alpha+\delta} = o(1), \text{ uniformly for } b > n^{\beta+\epsilon} \text{ as } n \nearrow \infty.$$

Here the convergence to 0 is justified because  $\alpha > 2$  and  $b > n^{\beta+\epsilon}$ . □

**Proof of Lemma 1.** From the definition of  $\Lambda_b(\cdot)$  and Lemma 8, we have:

$$\begin{aligned} \exp(\Lambda_b(\theta_{n,b})) &= \int_{-\infty}^b \exp(\theta_{n,b}x) F(dx) \\ &\leq 1 + c\theta_{n,b}^\kappa + e^{2\alpha} \bar{F}\left(\frac{2\alpha}{\theta_{n,b}}\right) + \exp(\theta_{n,b}) \bar{F}(b)(1 + o(1)), \end{aligned}$$

for  $\kappa$  as in (39). Usage of Lemma 8 is justified because  $\theta_n b = -\log(n\bar{F}(b)) \nearrow \infty$ . The last term,

$$\exp(\theta_{n,b}) \bar{F}(b) = \frac{1}{n\bar{F}(b)} \bar{F}(b) = \frac{1}{n}.$$

From Lemma 9, we have  $n\theta_{n,b}^\kappa = o(1)$  and  $\bar{F}(2\alpha/\theta_{n,b}) = o(1/n)$ , uniformly for  $b > n^{\beta+\epsilon}$ .

$$\text{Therefore, } \exp(\Lambda_b(\theta_n)) \leq 1 + \frac{1}{n}(1 + o(1)), \text{ as } n \nearrow \infty.$$

□

**Proof of Lemma 2** It is enough to show that, given  $\epsilon > 0$ , there exists  $b_\epsilon$  such that for all  $b > b_\epsilon$ ,

$$\inf_{k \geq 1} \frac{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\}}{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}} > 1 - \epsilon.$$

$$\begin{aligned} \mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\} &= \sum_{j=n_{k-1}+1}^{n_k} \mathbb{P}\{\tau = j, A_k\} \\ &\geq \sum_{j=n_{k-1}+1}^{n_k} \mathbb{P}\{\tau = j, S_i > -(A + i\delta) \text{ for all } i < j, X_j > b + A + j(\mu + \delta)\}, \end{aligned}$$

for some positive constants  $A$  and  $\delta$ . Let  $M_n := \max_{k \leq n} (S_k - k\mu)$  and  $M := \sup_k (S_k - k\mu)$ . Then,

$$\begin{aligned} \mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\} &\geq \sum_{j=n_{k-1}+1}^{n_k} \mathbb{P}\{M_{j-1} \leq b, S_i > -(A + i\delta) \text{ for all } i < j, X_j > b + A + j(\mu + \delta)\} \\ &\geq \sum_{j=n_{k-1}+1}^{n_k} \mathbb{P}\{M_{j-1} \leq b, S_i > -(A + i\delta) \text{ for all } i < j\} \bar{F}(b + A + j(\mu + \delta)) \\ &\geq \mathbb{P}\{M \leq b, S_i > -(A + i\delta) \text{ for all } i\} \sum_{j=n_{k-1}+1}^{n_k} \bar{F}(b + A + j(\mu + \delta)). \end{aligned}$$

Since  $\mathbb{P}\{M > b\} = o(1)$ , as  $b \nearrow \infty$ , by union bound, we have:

$$\mathbb{P}(\{M > b\} \cup \{S_i < -(A + i\delta) \text{ for some } i\}) \leq \epsilon + \mathbb{P}\{S_i < -(A + i\delta) \text{ for some } i\}.$$

Due to the law of large numbers, we can find  $i_\epsilon$  such that for all  $i > i_\epsilon$ ,  $S_i$  is larger than  $-(A + i\delta)$  with probability at least  $1 - \epsilon$ . Further for the collection  $(S_i : i \leq i_\epsilon)$ , we can choose  $A$  large enough such that for all  $i < i_\epsilon$ ,  $S_i$  is larger than  $-A$  with probability at least  $1 - \epsilon$ . Then,

$$\mathbb{P}(\{M > b\} \cup \{S_i < -(A + i\delta) \text{ for some } i\}) \leq 2\epsilon,$$

and hence,

$$\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\} \geq (1 - 2\epsilon) \sum_{j=n_{k-1}+1}^{n_k} \bar{F}(b + A + j(\mu + \delta)). \quad (40)$$

Now consider,

$$\begin{aligned} \sum_{j=n_{k-1}+1}^{n_k} \bar{F}(b + A + j(\mu + \delta)) &\geq \sum_{j=n_{k-1}+1}^{n_k} \int_j^{j+1} \bar{F}(b + A + u(\mu + \delta)) du \\ &\geq \int_{n_{k-1}+1}^{n_k} \bar{F}(b + A + u(\mu + \delta)) du. \end{aligned}$$

After changing the variables of integration, we get:

$$\begin{aligned} \sum_{j=n_{k-1}+1}^{n_k} \bar{F}(b + A + j(\mu + \delta)) &\geq \frac{1}{\mu + \delta} \int_{b+A+(n_{k-1}+1)(\mu+\delta)}^{b+A+n_k(\mu+\delta)} \bar{F}(u) du \\ &= \frac{\bar{F}_I(b + A + (n_{k-1} + 1)(\mu + \delta)) - \bar{F}_I(b + A + n_k(\mu + \delta))}{\mu + \delta}. \end{aligned}$$

Then from (40),

$$\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\} \geq \frac{1-2\epsilon}{\mu+\delta} \left( \bar{F}_I(b+A+(n_{k-1}+1)(\mu+\delta)) - \bar{F}_I(b+A+n_k(\mu+\delta)) \right).$$

Since  $\delta$  is arbitrary and  $\bar{F}_I(\cdot)$  is long-tailed, for values of  $b$  large enough, we have:

$$\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\} \geq \frac{(1-2\epsilon)(1-\epsilon)}{\mu} \left( \bar{F}_I(b+(n_{k-1}+1)\mu) - \bar{F}_I(b+n_k\mu) \right).$$

Now from (8), for all  $k$ ,

$$\mathbb{P}\{n_{k-1} < \tau_b \leq n_k, A_k\} \geq (1-2\epsilon)(1-\epsilon)^2 \mathbb{P}\{n_{k-1} < \tau_b \leq n_k\},$$

for large values of  $b$ , thus proving the claim.  $\square$

**Proof of Lemma 4.** Consider  $\theta : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ . From Lemma 8, we have that: for given  $\epsilon > 0$ , if  $x\theta(x) \nearrow \infty$ , then there exists  $x_\epsilon$  such that for all  $x > x_\epsilon$ ,

$$\int_{-\infty}^x e^{\theta(x)u} F(du) \leq 1 + c\theta^{1+\delta}(x) + e^{2\alpha} \bar{F}\left(\frac{2\alpha}{\theta(x)}\right) + e^{\theta(x)x} \bar{F}(x)(1+\epsilon),$$

for some  $\delta > 0$ . For this, we do not need any condition on left tail as in Assumption 1. By definition of  $\theta_k(b)$  in (20), we have  $(b+n_{k-1}\mu) \cdot \theta_k(b) \nearrow \infty$ , either if  $b$  or  $k$  grows to infinity. Expressing  $\theta_k(b)$  as  $\theta_k$ , for values of  $b$  and  $k$  satisfying  $b+n_{k-1}\mu > x_\epsilon$ , we have,

$$\begin{aligned} \exp(\Lambda_k(\theta_k)) &\leq 1 + c\theta_k^{1+\delta} + e^{2\alpha} \bar{F}\left(\frac{2\alpha}{\theta_k}\right) + e^{\theta_k \cdot (b+n_{k-1}\mu)} \bar{F}(b+n_{k-1}\mu)(1+\epsilon) \\ &\leq \exp\left(c\theta_k^{1+\delta} + e^{2\alpha} \bar{F}\left(\frac{2\alpha}{\theta_k}\right) + \frac{1}{n_k}(1+\epsilon)\right), \end{aligned}$$

because  $1+x \leq e^x$  and  $e^{\theta_k \cdot (b+n_{k-1}\mu)} \bar{F}(b+n_{k-1}\mu) = 1/n_k$ . Then,

$$\exp(n_k \Lambda_k(\theta_k)) \leq \exp\left(cn_k \theta_k^{1+\delta} + e^{2\alpha} n_k \bar{F}\left(\frac{2\alpha}{\theta_k}\right) + 1 + \epsilon\right). \quad (41)$$

Also see that,

$$n_k \theta_k^{1+\delta} = \frac{n_k}{(b+n_{k-1}\mu)^{1+\delta}} \left( \log\left(\frac{1}{n_k \bar{F}(b+n_{k-1}\mu)}\right) \right)^{1+\delta} < \epsilon, \quad (42)$$

if  $b$  and  $k$  are such that  $(b+n_{k-1}\mu)$  is large enough. Similarly for given  $\delta > 0$ , there exists  $x_\delta$  such that if  $b+n_{k-1}\mu > x_\delta$ , then

$$\frac{\bar{F}\left(\frac{2\alpha}{\theta_k}\right)}{\bar{F}(b+n_{k-1}\mu)} = \frac{\bar{F}\left(\frac{2\alpha(b+n_{k-1}\mu)}{-\log(n_k \bar{F}(b+n_{k-1}\mu))}\right)}{\bar{F}(b+n_{k-1}\mu)} \leq \left(\frac{1}{2\alpha} \log\left(\frac{1}{n_k \bar{F}(b+n_{k-1}\mu)}\right)\right)^{\alpha+\delta}.$$

Then for values of  $b$  and  $k$  such that  $(b+n_{k-1}\mu)$  is large enough,

$$\begin{aligned} n_k \bar{F}\left(\frac{2\alpha}{\theta_k}\right) &\leq n_k \bar{F}(b+n_{k-1}\mu) \left(\frac{1}{2\alpha} \log\left(\frac{1}{n_k \bar{F}(b+n_{k-1}\mu)}\right)\right)^{\alpha+\delta} \\ &= \frac{n_k L(b+n_{k-1}\mu)}{(b+n_{k-1}\mu)^\alpha} \left(\frac{1}{2\alpha} \log\left(\frac{1}{n_k \bar{F}(b+n_{k-1}\mu)}\right)\right)^{\alpha+\delta} < \epsilon, \end{aligned}$$

because  $\alpha > 2$ . Combining this with (41) and (42), for  $b$  and  $k$  such that  $b+n_{k-1}\mu$  is sufficiently large,

$$\exp(n_k \Lambda_k(\theta_k)) \leq \exp(1+3\epsilon),$$

thus establishing the claim.  $\square$

**Proof of Lemma 5.** From the uniform asymptotic (8), given  $\epsilon > 0$ , for  $b$  large enough and for any  $k \geq 1$ , we have:

$$\begin{aligned} \frac{n_k \bar{F}(b + n_{k-1}\mu)}{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}} &= \frac{n_k \bar{F}(b + n_{k-1}\mu)}{\frac{1}{\mu} \int_{b+n_{k-1}\mu}^{b+n_k\mu} \bar{F}(u) du} \cdot \frac{\frac{1}{\mu} \int_{b+n_{k-1}\mu}^{b+n_k\mu} \bar{F}(u) du}{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}} \\ &\leq (1 + \epsilon) \frac{n_k \bar{F}(b + n_{k-1}\mu)}{\frac{1}{\mu} \bar{F}(b + n_k\mu)(n_k\mu - n_{k-1}\mu)}. \end{aligned}$$

For  $k > 1, n_k = rn_{k-1}$ ; then,

$$\begin{aligned} \frac{n_k \bar{F}(b + n_{k-1}\mu)}{\mathbb{P}\{n_{k-1} < \tau_b \leq n_k\}} &\leq (1 + \epsilon) \frac{n_k \bar{F}(b + \frac{n_k}{r}\mu)}{\frac{n_k}{r} \bar{F}(b + n_k\mu)} \\ &\leq (1 + \epsilon) \frac{\bar{F}(\frac{1}{r}(b + n_k\mu))}{\frac{1}{r} \bar{F}(b + n_k\mu)} \\ &\leq (1 + \epsilon)^2 r^{\alpha+1}, \end{aligned}$$

for values of  $b$  large enough, and for any  $k \geq 1$ , because of the regularly varying nature of  $\bar{F}(\cdot)$ . Also,

$$\frac{n_1 \bar{F}(b + n_0\mu)}{\mathbb{P}\{n_0 < \tau_b \leq n_1\}} \leq (1 + \epsilon) \frac{r \bar{F}(b)}{r \bar{F}(b + r\mu)} \leq (1 + \epsilon)^2,$$

for large values of  $b$ , because  $F(\cdot)$  is long-tailed. Thus for any  $k$ , we can find a constant  $c_2$  such that the claim holds.  $\square$